

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 16:59:33 ; Search time 30893 Seconds
(without alignments)
11720.704 Million cell updates/sec

Title: US-10-029-020-13
Perfect score: 8354
Sequence: 1 gtttggtgagtgaggagc.....ccggagggtgacagaggac 8354

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htg.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8354	100.0	8354	6	AX556500 Sequence
2	8225	98.5	8645	6	AX600210 Sequence
3	7779.2	93.1	8438	6	AX675551 Sequence
4	6975.4	83.5	8585	10	AB025413 Mus muscu
5	6601.2	79.0	9722	10	AF059485 Mus muscu
6	4422.8	52.9	9264	5	AB026980 Dario rer
7	4197.2	50.2	8993	9	BSM806812 Homo sapi
8	3680.8	44.1	8624	9	AB037723 Homo sapi
9	3544.6	42.4	8964	10	AB025412 Mus muscu
10	3502.2	41.9	8645	6	AX623355 Sequence
11	3473.8	41.6	8675	6	AX662353 Sequence
12	3422	41.0	8473	6	AX662357 Sequence
13	3398.8	40.7	8487	6	AX662359 Sequence
14	3397.8	40.7	8689	6	AX250067 Sequence
15	3397.8	40.7	8689	10	AF086607 Rattus no
16	3385	40.5	9266	6	AX250008 Sequence
17	3384.8	40.5	8797	6	AX250063 Sequence
18	3384.8	40.5	8797	6	AX250066 Sequence
19	3384.8	40.5	8797	10	AB025411 Mus muscu
20	3351.4	40.1	9729	6	AX250013 Sequence
21	3330.4	39.9	8373	10	AB025410 Mus muscu
22	3320	39.7	8118	5	GGA236613 Gallus ga
23	3280	39.3	7816	10	AF195418 Mus muscu
24	3258	39.0	8297	9	AF100772 Homo sapi
25	3219	38.5	3486	9	BSM806114 Homo sapi
26	3165.8	37.9	8816	5	AB026979 Dario rer
27	3158.4	37.8	8575	6	AX921803 Sequence
28	3137.4	37.6	7781	9	AB032953 Homo sapi
29	3070.8	36.8	8409	5	GA2279031 Gallus ga
30	3070.8	36.8	8409	6	AX250068 Sequence
31	3055.6	36.6	3320	9	AK127101 Homo sapi
32	3051.8	36.5	5583	10	AK122490 Mus muscu
33	2925.2	35.0	7713	9	BSM808325 Homo sapi
34	2794.6	33.5	3048	9	AK127705 Homo sapi
35	2622.2	31.4	5804	10	AK122513 Mus muscu
36	2533	30.2	6560	6	AX250065 Sequence
37	2463	29.5	5309	9	AB040888 Homo sapi
38	2437.6	29.2	2685	9	AK056531 Homo sapi
39	1708.8	20.5	3111	6	BD270887 Method of
40	1708.8	20.5	3111	6	AR270934 Sequence
41	1708.8	20.5	3111	6	BD140153 Utilizati
42	1708.8	20.5	3111	9	AF009227 Homo sapi
43	1673.6	20.0	1680	6	AR270935 Sequence
44	1618.8	19.4	170156	2	AP002412 Homo sapi
45	1618.8	19.4	186084	9	AP002768 Homo sapi

ALIGNMENTS

RESULT 1
AX556500
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

AX556500
Sequence
AX556500
AX556500.1

8354 bp
from Patent WO02057453.
GI:25899736

DNA
linear

PAT 27-NOV-2002

Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Gangolli, E.A., Patturajan, M., Vernet, C.A., Malyankar, U.M.,
Kekuda, R., Stone, D.J., Anderson, D., Shimkets, R.A., Burgess, C.E.,
Zerhusen, B.D., Liu, X., Spytek, K.A., Casman, S.J., Boldog, F.L.,

Smithson, G., Li, L., and Ji, W.
 Polypeptides and nucleic acids encoding same
 Patent: WO 02057453-A 13 25-JUL-2002;
 Curagen Corporation (US)
 Location/Qualifiers
 1..8354
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

source

ORIGIN

Query Match 100.0%; Score 8354; DB 6; Length 8354;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 8354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTTGTGATGGAGGAGCGGGCCGAGGCCATGACGTGAAGGAGAGAGCCTTA 60
 DB 1 GTTGTGATGGAGGAGCGGGCCGAGGCCATGACGTGAAGGAGAGAGCCTTA 60

QY 61 CGCTCGCTGACCCCGGCGCGGAGCGCGGCGGCTACACAGCTCGTCCGCGGACAG 120
 DB 61 CGCTCGCTGACCCCGGCGCGGAGCGCGGCGGCTACACAGCTCGTCCGCGGACAG 120

QY 121 CGAGAGGGCAAGCCCGCGAGAAATGTAAGCTCAGCGAGACCCCTGAAGGCTACGA 180
 DB 121 CGAGAGGGCAAGCCCGCGAGAAATGTAAGCTCAGCGAGACCCCTGAAGGCTACGA 180

QY 181 CCAGAGCCCGGCTGAGCTATGGCAGCGCGGTCAAGGACATTTGCGCGAGGAGCGGA 240
 DB 181 CCAGAGCCCGGCTGAGCTATGGCAGCGCGGTCAAGGACATTTGCGCGAGGAGCGGA 240

QY 241 GGAATTTCCGCGACAGGTGCAACTTCAACCTCCGCGAGCTGGGGTGGAGAGTAAC 300
 DB 241 GGAATTTCCGCGACAGGTGCAACTTCAACCTCCGCGAGCTGGGGTGGAGAGTAAC 300

QY 301 GCCCCTCAGCGGACCTGTACCGGACAGATTTGGCTGCCCAATGCCGCTACTCCAT 360
 DB 301 GCCCCTCAGCGGACCTGTACCGGACAGATTTGGCTGCCCAATGCCGCTACTCCAT 360

QY 361 GGGGCTGGCTCTGATGCCGACATGGAGGCTGACAGCGTGTCTCCCTGAGCACCCTGT 420
 DB 361 GGGGCTGGCTCTGATGCCGACATGGAGGCTGACAGCGTGTCTCCCTGAGCACCCTGT 420

QY 421 CGCTGTGGGCGGAGACACCGTCAAGGCGAGCTCTGCTGTCTGTCAGCGCGGCA 480
 DB 421 CGCTGTGGGCGGAGACACCGTCAAGGCGAGCTCTGCTGTCTGTCAGCGCGGCA 480

QY 481 TTCCAATCTCACACTCACCGACACCGGACATGAAACACATGAGACTGATCATCGGGCGG 540
 DB 481 TTCCAATCTCACACTCACCGACACCGGACATGAAACACATGAGACTGATCATCGGGCGG 540

QY 541 CCTGAGAACACCGGGGGCTCCGAGCGCGCGCGCGCTCTGCGACGCCACACCCC 600
 DB 541 CCTGAGAACACCGGGGGCTCCGAGCGCGCGCGCGCTCTGCGACGCCACACCCC 600

QY 601 CAACAGCACACCGGCTCTCACTTCACTCCTGAACCGGGGCACTTACGCGGAGGAG 660
 DB 601 CAACAGCACACCGGCTCTCACTTCACTCCTGAACCGGGGCACTTACGCGGAGGAG 660

QY 661 CAACCCAGCCGCGCGCCACCGACCACTCGCTCTCCGAGAGCCCTCCCGCGGCGC 720
 DB 661 CAACCCAGCCGCGCGCCACCGACCACTCGCTCTCCGAGAGCCCTCCCGCGGCGC 720

QY 721 CCAGAGCCTCCACCGCCAGGAGACTGGCTCTCAACAGCAACATCCCTCTGAGAC 780
 DB 721 CCAGAGCCTCCACCGCCAGGAGACTGGCTCTCAACAGCAACATCCCTCTGAGAC 780

QY 781 CAGGAACCTAGGCAAGCAGCATTCTTAGGACATTTGAGGACAACTCTATTGAGATGA 840
 DB 781 CAGGAACCTAGGCAAGCAGCATTCTTAGGACATTTGAGGACAACTCTATTGAGATGA 840

QY 841 CATTTCTGGGCTCCCGGCTATGGGCTTACAGTACGGGCACTTCTCTTCAAGCC 900
 DB 841 CATTTCTGGGCTCCCGGCTATGGGCTTACAGTACGGGCACTTCTCTTCAAGCC 900

841 CATTTCTGGGCTCCCGGCTATGGGCTTACAGTACGGGCACTTCTCTTCAAGCC 900
 QY 901 TGGAGGACCTCCCGCTCTTCTGACACACATCACAGGGTACCCACTGACGCTCCAGAC 960
 DB 901 TGGAGGACCTCCCGCTCTTCTGACACACATCACAGGGTACCCACTGACGCTCCAGAC 960

QY 961 AGTGTACTCTCTCCGCGCGGACCCCTGCGCGGAGACCTTTCGCGCGGCGGCTTAA 1020
 DB 961 AGTGTACTCTCTCCGCGCGGACCCCTGCGCGGAGACCTTTCGCGCGGCGGCTTAA 1020

QY 1021 CTTCAAGAGCCCTCCAAGTACTGTAAGTGGAGTGCAGCCCTGAGCGCATCTCAT 1080
 DB 1021 CTTCAAGAGCCCTCCAAGTACTGTAAGTGGAGTGCAGCCCTGAGCGCATCTCAT 1080

QY 1081 CTCAGCCACTCTGTCATCTCTGTCATCTTTGTGGCATGACCTTTTGGCCTTAA 1140
 DB 1081 CTCAGCCACTCTGTCATCTCTGTCATCTTTGTGGCATGACCTTTTGGCCTTAA 1140

QY 1141 CTGGCACCTGACAGCCGATGGAGGGCAGATGTATGAGATCAGGAGGACACAGCCAG 1200
 DB 1141 CTGGCACCTGACAGCCGATGGAGGGCAGATGTATGAGATCAGGAGGACACAGCCAG 1200

QY 1201 TTGGCTGTGCAACCGAGCTCTCCTATACCCCTCAGGGGCACTGGCTTAGAGACCC 1260
 DB 1201 TTGGCTGTGCAACCGAGCTCTCCTATACCCCTCAGGGGCACTGGCTTAGAGACCC 1260

QY 1261 TGACAGGAAAGCAAGAAACACAGAAAGAAAGCCAGTAGTTCTTTCCAGAGGACAG 1320
 DB 1261 TGACAGGAAAGCAAGAAACACAGAAAGAAAGCCAGTAGTTCTTTCCAGAGGACAG 1320

QY 1321 TTTCTATAGATTCTGAGAAATTTGATGTGGAAAGGCGAGCTCCAGAAAGATTCCTCTGG 1380
 DB 1321 TTTCTATAGATTCTGAGAAATTTGATGTGGAAAGGCGAGCTCCAGAAAGATTCCTCTGG 1380

QY 1381 CACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGTGATCTGAAATCAATGTGC 1440
 DB 1381 CACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGTGATCTGAAATCAATGTGC 1440

QY 1441 TCTGGAAAGGCGAGCCCTGGTTGGCATTTATGGCAGAAAGGCGCTCCCTCTTCAATAC 1500
 DB 1441 TCTGGAAAGGCGAGCCCTGGTTGGCATTTATGGCAGAAAGGCGCTCCCTCTTCAATAC 1500

QY 1501 ACAGTTTGAATTTGTGGAGCTCTGGATGGCAGAGGCTCTTAACCCAGAGGCGGAG 1560
 DB 1501 ACAGTTTGAATTTGTGGAGCTCTGGATGGCAGAGGCTCTTAACCCAGAGGCGGAG 1560

QY 1561 CTTAGAGGGGACCCCGCGGCACTCTCGGGGAACTGTGCCCCCTCCAGCCATGAGACAGG 1620
 DB 1561 CTTAGAGGGGACCCCGCGGCACTCTCGGGGAACTGTGCCCCCTCCAGCCATGAGACAGG 1620

QY 1621 CTTCTATCCAGTATTTGGATTCAGGAATCTGGCATCTGGCTTTTACAAATGACGAAAGGA 1680
 DB 1621 CTTCTATCCAGTATTTGGATTCAGGAATCTGGCATCTGGCTTTTACAAATGACGAAAGGA 1680

QY 1681 GTCAAGTGGTTTCTTCTCACCACTGCCATTTGAGTCGGTGAATTAATGACGAAAGGA 1740
 DB 1681 GTCAAGTGGTTTCTTCTCACCACTGCCATTTGAGTCGGTGAATTAATGACGAAAGGA 1740

QY 1741 CTGCTATGGCAATTTGATCTCTCTGGAGCTTGCATCTGCTTCTTCTGGGTTTCTGGG 1800
 DB 1741 CTGCTATGGCAATTTGATCTCTCTGGAGCTTGCATCTGCTTCTTCTGGGTTTCTGGG 1800

QY 1801 CCCCAGCTGTGGCAGAGCTCTCTCCCGTCTCTGTAGCGGAATGGCCAAATCATGAA 1860
 DB 1801 CCCCAGCTGTGGCAGAGCTCTCTCCCGTCTCTGTAGCGGAATGGCCAAATCATGAA 1860

QY 1861 AGGCAGATGCTTTGTGCCACAGTGGCTGGAAGGCGCTGAGTGCATGTGCCACCAACCA 1920
 DB 1861 AGGCAGATGCTTTGTGCCACAGTGGCTGGAAGGCGCTGAGTGCATGTGCCACCAACCA 1920

QY 1921 GTGTATCATGTGGCTTGCAGAACCACTTGCATCATCGGGCACCTCTCTTCAAGCC 1980
 DB 1921 GTGTATCATGTGGCTTGCAGAACCACTTGCATCATCGGGCACCTCTCTTCAAGCC 1980

QY	1981	CCCTGGCTACAAAGGCGAGAGCTGTGAGAAAGTGCATGCAATGACCCACATGTTGAGG	2040
Db	1981	CCCTGGCTACAAAGGCGAGAGCTGTGAGAAAGTGCATGCAATGACCCACATGTTGAGG	2040
QY	2041	CCGGGGTGTGCTGTGAGAGGCAATGCAATGCTTTGTGGATGGGAGGACCAAACTG	2100
Db	2041	CCGGGGTGTGCTGTGAGAGGCAATGCAATGCTTTGTGGATGGGAGGACCAAACTG	2100
QY	2101	CGAGACCCCGAGGCGACATGCTTAGACCAAGTGTTCAGGCGACAGCACTTCCTCCGGG	2160
Db	2101	CGAGACCCCGAGGCGACATGCTTAGACCAAGTGTTCAGGCGACAGCACTTCCTCCGGG	2160
QY	2161	CACCGGGTTTGACCTGTGACCCCAAGCTGACATGACACCACTGTTCTATCGAGATCTG	2220
Db	2161	CACCGGGTTTGACCTGTGACCCCAAGCTGACATGACACCACTGTTCTATCGAGATCTG	2220
QY	2221	TGCTGCCACTGTGCTGCCATGCGTGTGCTAGGAGGCGACCTGCCCTCGAGGATGG	2280
Db	2221	TGCTGCCACTGTGCTGCCATGCGTGTGCTAGGAGGCGACCTGCCCTCGAGGATGG	2280
QY	2281	CTGATGGGGCAGCTCGACACAGCGGGCTGCGACCCGGCTGTGCCAGCATGGGAC	2340
Db	2281	CTGATGGGGCAGCTCGACACAGCGGGCTGCGACCCGGCTGTGCCAGCATGGGAC	2340
QY	2341	CTGCCGCGAGGCAAGTGGAGTGCAGCCCTGGCTGGAAATGGCGAACTGCACCATCGC	2400
Db	2341	CTGCCGCGAGGCAAGTGGAGTGCAGCCCTGGCTGGAAATGGCGAACTGCACCATCGC	2400
QY	2401	TCACATCTGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTCAATGCAACGGCAG	2460
Db	2401	TCACATCTGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTCAATGCAACGGCAG	2460
QY	2461	ATGTACCTTAGACCTGAAATGTTGGCACTGGCTGTGCCAGTGGGCTGGAGAGTGG	2520
Db	2461	ATGTACCTTAGACCTGAAATGTTGGCACTGGCTGTGCCAGTGGGCTGGAGAGTGG	2520
QY	2521	CTGTGACACTTCATGAGACTGCTGCGGTGACAGCAAGCAATGATGAGATGGCCT	2580
Db	2521	CTGTGACACTTCATGAGACTGCTGCGGTGACAGCAAGCAATGATGAGATGGCCT	2580
QY	2581	GGTGACTGCATGGACCCCTGACTGCTGCTCAGCCCTGTGCCATATCAACCCGCTGTG	2640
Db	2581	GGTGACTGCATGGACCCCTGACTGCTGCTCAGCCCTGTGCCATATCAACCCGCTGTG	2640
QY	2641	CCTTGGCTCCCTTAACCTCTGGACATCATCAGAGACACAGGTCCCTGTGTACAGCA	2700
Db	2641	CCTTGGCTCCCTTAACCTCTGGACATCATCAGAGACACAGGTCCCTGTGTACAGCA	2700
QY	2701	GAACCTACACTCCTTCTATGACCGCATCAAGTTCCTGCTGGCAGGACAGCAGCAT	2760
Db	2701	GAACCTACACTCCTTCTATGACCGCATCAAGTTCCTGCTGGCAGGACAGCAGCAT	2760
QY	2761	AATCCCGGGGAGAACCCCTTTGATGGAGGCGATGCTTGTGTATTCGTGCCAAAGTGT	2820
Db	2761	AATCCCGGGGAGAACCCCTTTGATGGAGGCGATGCTTGTGTATTCGTGCCAAAGTGT	2820
QY	2821	GACATCAGATGGAACCCCTGTTGTGTGAAATCATCAGTTTGTCAATACCCCTCTCT	2880
Db	2821	GACATCAGATGGAACCCCTGTTGTGTGAAATCATCAGTTTGTCAATACCCCTCTCT	2880
QY	2881	TGGATATACATCAGCAGCAGATGGCAGCTTTGACCTTGTGACAAATGCGCGCATCTC	2940
Db	2881	TGGATATACATCAGCAGCAGATGGCAGCTTTGACCTTGTGACAAATGCGCGCATCTC	2940
QY	2941	CATCATCTGGGTTGAGAGGGGACCTTTGATCAGAGAGACACCTGTGGCTGCC	3000
Db	2941	CATCATCTGGGTTGAGAGGGGACCTTTGATCAGAGAGACACCTGTGGCTGCC	3000
QY	3001	ATGGGATCGCTTCTTGTGATGGAACCATCATCATGAGACATGAGGAGATGAGATTCC	3060
Db	3001	ATGGGATCGCTTCTTGTGATGGAACCATCATCATGAGACATGAGGAGATGAGATTCC	3060
QY	3061	CAGCTGTGACCTGAGCAATTTTGGCGCCGCCAACCCAGCTGCTCTCCATCCCACTGAC	3120
Db	3061	CAGCTGTGACCTGAGCAATTTTGGCGCCGCCAACCCAGCTGCTCTCCATCCCACTGAC	3120
QY	3121	GTCTTTGGCAGCTCTGTGAGAGAAAGCCCATTTGTCCGGAATTCAGGCTTTGCA	3180
Db	3121	GTCTTTGGCAGCTCTGTGAGAGAAAGCCCATTTGTCCGGAATTCAGGCTTTGCA	3180
QY	3181	GGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCGGACCCC	3240
Db	3181	GGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCGGACCCC	3240
QY	3241	TGCTCAAAATCTGTCTGAGGATCAGCTCACCCACCGACCATCCCTTCAACCTCAT	3300
Db	3241	TGCTCAAAATCTGTCTGAGGATCAGCTCACCCACCGACCATCCCTTCAACCTCAT	3300
QY	3301	GAAGGTGCACCTCATGATAGCGGTGAGGCGCGCTTTCAGGAAGTGTGCTGCGACG	3360
Db	3301	GAAGGTGCACCTCATGATAGCGGTGAGGCGCGCTTTCAGGAAGTGTGCTGCGACG	3360
QY	3361	CCGAGACCTGTCTTATTTTCAATTTGGACAAGCAGACGCTCTACAACCCAGAAGTGT	3420
Db	3361	CCGAGACCTGTCTTATTTTCAATTTGGACAAGCAGACGCTCTACAACCCAGAAGTGT	3420
QY	3421	TGGGCTTTCAGAAAGCTTTTTCCTGGGTTATGAATATGAATCTGCGCCAGATCTAAT	3480
Db	3421	TGGGCTTTCAGAAAGCTTTTTCCTGGGTTATGAATATGAATCTGCGCCAGATCTAAT	3480
QY	3481	CCTGTGGGAAAAAGAACCAACAGTGTGAGGCTATGAAATTTGACCGCTCAAGCTTG	3540
Db	3481	CCTGTGGGAAAAAGAACCAACAGTGTGAGGCTATGAAATTTGACCGCTCAAGCTTG	3540
QY	3541	AGGATGAGCCTAGACAAACATCATGCTTCAACATTTCAAAGTGTGATCTGTCACAAAG	3600
Db	3541	AGGATGAGCCTAGACAAACATCATGCTTCAACATTTCAAAGTGTGATCTGTCACAAAG	3600
QY	3601	GAATGGGGAACCAAGTTTGTGCTCAGCAGCCTCTCTCATTTGGGAGCATCATGGCAA	3660
Db	3601	GAATGGGGAACCAAGTTTGTGCTCAGCAGCCTCTCTCATTTGGGAGCATCATGGCAA	3660
QY	3661	TGGGCGCGGAGAGCATCTCTGCCCGCAGCTGCAACGGCTTGTGACGGCAACAGCT	3720
Db	3661	TGGGCGCGGAGAGCATCTCTGCCCGCAGCTGCAACGGCTTGTGACGGCAACAGCT	3720
QY	3721	CCTGGCGCCAGTGGCCCTCACCTGGCTGTGACGGGAGCTCTATGTTGGTATTCAA	3780
Db	3721	CCTGGCGCCAGTGGCCCTCACCTGGCTGTGACGGGAGCTCTATGTTGGTATTCAA	3780
QY	3781	CTACATTAGAAGGATCTTCCCTCTGGAATGTCAACCAATCTCTAGAGCTGAGGAATAA	3840
Db	3781	CTACATTAGAAGGATCTTCCCTCTGGAATGTCAACCAATCTCTAGAGCTGAGGAATAA	3840
QY	3841	AGATTTAGACATAGTACAGTCCAGACACAAATATCTACCTGGCCACAGACCCCATGAG	3900
Db	3841	AGATTTAGACATAGTACAGTCCAGACACAAATATCTACCTGGCCACAGACCCCATGAG	3900
QY	3901	TGGGCGCGCTCTCTCTTTGACAGCAACAGCGCGGCTTTTAAATCAAGTCCACTGT	3960
Db	3901	TGGGCGCGCTCTCTCTTTGACAGCAACAGCGCGGCTTTTAAATCAAGTCCACTGT	3960
QY	3961	GGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGTTGGGGGACAGGTGAACAGTGCCT	4020
Db	3961	GGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGTTGGGGGACAGGTGAACAGTGCCT	4020
QY	4021	CCCTTTGATGACATCGCTGGGGATGGTGGAGGGCCACAGAGCCACACTCACCAA	4080
Db	4021	CCCTTTGATGACATCGCTGGGGATGGTGGAGGGCCACAGAGCCACACTCACCAA	4080
QY	4081	TCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCTGGTGGACCATGAT	4140
Db	4081	TCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCTGGTGGACCATGAT	4140
QY	4141	CAGACGATCGATCAAGATGGGATCATCTCCACCCCTGCTCGGCTCTAATGATCTCATC	4200

Db 4141 CAGAGCATCGATCAGATGGGATCATCTCCACCCCTGCTCGGCTCTAATGATCTCATCTC 4200
Qy |||||
4201 AGCCGGGCACCTACGCTGTGATTCCTGATGATATTTCCAGGTAAAGATGGAGTGGCC 4260
Db |||||
4201 AGCCGGGCACCTACGCTGTGATTCCTGATGATATTTCCAGGTAAAGATGGAGTGGCC 4260
Qy |||||
4261 CACAGACTTAGCCATCAACCAATGGAACAACTCACTTTATCTCTCCGACACATATGGT 4320
Db |||||
4261 CACAGACTTAGCCATCAACCAATGGAACAACTCACTTTATCTCTCCGACACATATGGT 4320
Qy |||||
4321 CCTGAAATCTCTGAAATCCACAGGTGCGCAATTTGCGCGGGAGGCCCATGCACTGCCA 4380
Db |||||
4321 CCTGAAATCTCTGAAATCCACAGGTGCGCAATTTGCGCGGGAGGCCCATGCACTGCCA 4380
Qy |||||
4381 GGTCCCTGGCATTTGACCACTTCCTGCTAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTC 4440
Db |||||
4381 GGTCCCTGGCATTTGACCACTTCCTGCTAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTC 4440
Qy |||||
4441 AGCCACCGCTTTGGCTGTTTACACAAATGGGGTCTCTGATATTTGCTGAGACTGATGAAA 4500
Db |||||
4441 AGCCACCGCTTTGGCTGTTTACACAAATGGGGTCTCTGATATTTGCTGAGACTGATGAAA 4500
Qy |||||
4501 AAGAGTCAACCGCATCAGCGAGGTCAACACTAGTGGAGAGATCTCACTGTTGCTGGGC 4560
Db |||||
4501 AAGAGTCAACCGCATCAGCGAGGTCAACACTAGTGGAGAGATCTCACTGTTGCTGGGC 4560
Qy |||||
4561 CCCAGTGGCTGTGACTGTGAAATGATGCCAACTGTGATGTTTCTGGAGACGATGG 4620
Db |||||
4561 CCCAGTGGCTGTGACTGTGAAATGATGCCAACTGTGATGTTTCTGGAGACGATGG 4620
Qy |||||
4621 TTATGCCAAGATGCAAAAGTTAAATPACCCCATCTTCCTTGGCTGTGTGCTGATGGGA 4680
Db |||||
4621 TTATGCCAAGATGCAAAAGTTAAATPACCCCATCTTCCTTGGCTGTGTGCTGATGGGA 4680
Qy |||||
4681 GCTCTACCTGCGGACCTTGGACATCCGAAATTCGGTTTATCCGGAAGACAGCCCTTT 4740
Db |||||
4681 GCTCTACCTGCGGACCTTGGACATCCGAAATTCGGTTTATCCGGAAGACAGCCCTTT 4740
Qy |||||
4741 CCTCAACACCCAGAACATGTAGCTGTCTTACCAATTCACGAGGAGCTCTATCTGTT 4800
Db |||||
4741 CCTCAACACCCAGAACATGTAGCTGTCTTACCAATTCACGAGGAGCTCTATCTGTT 4800
Qy |||||
4801 TGATACCAACCGCAAGCACTGTATACCCAAAGCCTGCCACAGGAGACTACCTGTACAA 4860
Db |||||
4801 TGATACCAACCGCAAGCACTGTATACCCAAAGCCTGCCACAGGAGACTACCTGTACAA 4860
Qy |||||
4861 CTTCACTTACACTGGGGACGGGACATCACACTCATCACAGACAAACATGGCAACATGGT 4920
Db |||||
4861 CTTCACTTACACTGGGGACGGGACATCACACTCATCACAGACAAACATGGCAACATGGT 4920
Qy |||||
4921 AAATGTCCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTGTCCAGATGCCAGGT 4980
Db |||||
4921 AAATGTCCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTGTCCAGATGCCAGGT 4980
Qy |||||
4981 GTACTGGGTGACCATGGGACCAACAGTGCATCAAGAGTGTGACACACAAAGACACGA 5040
Db |||||
4981 GTACTGGGTGACCATGGGACCAACAGTGCATCAAGAGTGTGACACACAAAGACACGA 5040
Qy |||||
5041 GTTGCCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAA 5100
Db |||||
5041 GTTGCCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAA 5100
Qy |||||
5101 CGGATGGACACATTTATGATGATGACAGAGTTGGCGCTGACAAATGTGACCTTCCC 5160
Db |||||
5101 CGGATGGACACATTTATGATGATGACAGAGTTGGCGCTGACAAATGTGACCTTCCC 5160
Qy |||||
5161 TACTGGCCAGGTGAGCAGTTTCCGAAGTATACAGACAGTTTCACTGATGTCCAGGTAGA 5220
Db |||||
5161 TACTGGCCAGGTGAGCAGTTTCCGAAGTATACAGACAGTTTCACTGATGTCCAGGTAGA 5220
Qy |||||
5221 GACCTCAGCAAGATGATGTACCAATTAACCAACACTGTCTGCTCAGCGGCTTCTA 5280

Db 5221 GACCTCAGCAAGATGATGTACCATACCAACCACTGTCTGCTCAGCGGCTTCTA 5280
Qy |||||
5281 CACACTGTCTGAAGACCAAGTCCGGAAACAGCTACTACTATCGGGGCGGATGGCTCTTGGC 5340
Db |||||
5281 CACACTGTCTGAAGACCAAGTCCGGAAACAGCTACTACTATCGGGGCGGATGGCTCTTGGC 5340
Qy |||||
5341 GTTGTCTGTGGCCAAAGCGCATGAGGTGGCGCTGAGACTGAGCCCACTTGTCTGGCTGG 5400
Db |||||
5341 GTTGTCTGTGGCCAAAGCGCATGAGGTGGCGCTGAGACTGAGCCCACTTGTCTGGCTGG 5400
Qy |||||
5401 CACCGTCAACCCCAACCGTGGCAAGAGAAATGTACGCTGGCCATTCGACAAACGGGCTCAA 5460
Db |||||
5401 CACCGTCAACCCCAACCGTGGCAAGAGAAATGTACGCTGGCCATTCGACAAACGGGCTCAA 5460
Qy |||||
5461 CTTGTGTGAGTGGCGCCAGCGCAAGAGACAGCTCGGGGCCAGGTCATGTCTTTTGGGCG 5520
Db |||||
5461 CTTGTGTGAGTGGCGCCAGCGCAAGAGACAGCTCGGGGCCAGGTCATGTCTTTTGGGCG 5520
Qy |||||
5521 CCGGCTGCGGGTGCACAAACCGAAATCTCTATCTCTGGACTTTGATCGCGTAACACGCAC 5580
Db |||||
5521 CCGGCTGCGGGTGCACAAACCGAAATCTCTATCTCTGGACTTTGATCGCGTAACACGCAC 5580
Qy |||||
5581 AGAAGATCTATGATGACACCGCAAGTTCAACCTTCGGATTCTGTAGCACCGGCGG 5640
Db |||||
5581 AGAAGATCTATGATGACACCGCAAGTTCAACCTTCGGATTCTGTAGCACCGGCGG 5640
Qy |||||
5641 GCGGCCACGCTCTGGTCACCCAGCAGGCTGAATGGTGTCAACGTGACATCTCCCC 5700
Db |||||
5641 GCGGCCACGCTCTGGTCACCCAGCAGGCTGAATGGTGTCAACGTGACATCTCCCC 5700
Qy |||||
5701 TGGGGTTTACATTTGCTGGCATCCAGGGGCAATGCTGAAAGAAATGGAATACGACCA 5760
Db |||||
5701 TGGGGTTTACATTTGCTGGCATCCAGAGGGCATCATGCTGTGAAAGAAATGGAATACGACCA 5760
Qy |||||
5761 GCGGGCCGATCAGATCCAGGATCTTCTGTATGGGAAGACATGAGCTACACATCTT 5820
Db |||||
5761 GCGGGCCGATCAGATCCAGGATCTTCTGTATGGGAAGACATGAGCTACACATCTT 5820
Qy |||||
5821 AGAAGATCCATGGTGTCTACTACAGCCAGAGGAGTATATCTTTTGGATTGCAAA 5880
Db |||||
5821 AGAAGATCCATGGTGTCTACTACAGCCAGAGGAGTATATCTTTTGGATTGCAAA 5880
Qy |||||
5881 GATGACCGCTCTCTTCTGTGACCATGCCAAACGTCGGCGGACAGACACTAGAGACCAT 5940
Db |||||
5881 GATGACCGCTCTCTTCTGTGACCATGCCAAACGTCGGCGGACAGACACTAGAGACCAT 5940
Qy |||||
5941 CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGCAATGCTCAGTCTAT 6000
Db |||||
5941 CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGCAATGCTCAGTCTAT 6000
Qy |||||
6001 ACAGACTTCACTGAGGATGGGACCTCTTCAACACTTCTACCTGGGCACTGGCCGAG 6060
Db |||||
6001 ACAGACTTCACTGAGGATGGGACCTCTTCAACACTTCTACCTGGGCACTGGCCGAG 6060
Qy |||||
6061 GGTGATATACAGTATGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAAA 6120
Db |||||
6061 GGTGATATACAGTATGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAAA 6120
Qy |||||
6121 GGTGATATACAGTATGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAAA 6180
Db |||||
6121 GGTGATATACAGTATGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAAA 6180
Qy |||||
6181 GGGCTTCACTGACCATCCCTACCGTCAGATTGGGCCCCCTGATTGACCGACAGATCTT 6240
Db |||||
6181 GGGCTTCACTGACCATCCCTACCGTCAGATTGGGCCCCCTGATTGACCGACAGATCTT 6240
Qy |||||
6241 CCGCTTCACTGAGGAAGCATGGTCAACGCCCGTTTGTACTACAACTATATGACACAGCTT 6300
Db |||||
6241 CCGCTTCACTGAGGAAGCATGGTCAACGCCCGTTTGTACTACAACTATATGACACAGCTT 6300
Qy |||||
6301 CCGGTTGACAGCATGAGGCTGTGATCAAGAGACCCCACTGCCATTGTCTTATCG 6360
Db |||||
6301 CCGGTTGACAGCATGAGGCTGTGATCAAGAGACCCCACTGCCATTGTCTTATCG 6360

QY	6361	CTATGATGATGTGTGAGCAAGACAGAGCAGTGTGGGAAGTTGGTGTCTATTACTATGA	6420
Db	6361	CTATGATGATGTGTGAGCAAGACAGAGCAGTGTGGGAAGTTGGTGTCTATTACTATGA	6420
QY	6421	CATTAAACAGATCATCAACACAGCTGTATGACCAACACCAAGCATTTTGTATGCATATGG	6480
Db	6421	CATTAAACAGATCATCAACACAGCTGTATGACCAACACCAAGCATTTTGTATGCATATGG	6480
QY	6481	CAGGATGAAGAAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTGATGACCGTCCA	6540
Db	6481	CAGGATGAAGAAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTGATGACCGTCCA	6540
QY	6541	GTATGATTAACATGGGGGAGTAGTAGAAGAGAGCTGAAGTAGGACCTTACGCCAATAC	6600
Db	6541	GTATGATTAACATGGGGGAGTAGTAGAAGAGAGCTGAAGTAGGACCTTACGCCAATAC	6600
QY	6601	CACCTCGCTACTCTATGATGATGATGCTGACGGCAGCTGCAGACAGTCTCCATCAATGA	6660
Db	6601	CACCTCGCTACTCTATGATGATGATGCTGACGGCAGCTGCAGACAGTCTCCATCAATGA	6660
QY	6661	CAAGCCACTCTGGCGCTACAGCTACGACTACGACTCAATGGGAACCTGACCTTCTAGCCCTGG	6720
Db	6661	CAAGCCACTCTGGCGCTACAGCTACGACTACGACTCAATGGGAACCTGACCTTCTAGCCCTGG	6720
QY	6721	GACAGTGCACGGCTCAACACACTACGGTATGACATCCGGACCGCATCACTCGGCTGGG	6780
Db	6721	GACAGTGCACGGCTCAACACACTACGGTATGACATCCGGACCGCATCACTCGGCTGGG	6780
QY	6781	TGACGTGCAATACAAAGTGGATGAGTGGCTTCTGAGGCAGCGGGCGGTGATATCTT	6840
Db	6781	TGACGTGCAATACAAAGTGGATGAGTGGCTTCTGAGGCAGCGGGCGGTGATATCTT	6840
QY	6841	TGAGTACAACTCAGTGGCTGCTCATCAAGCCCTACAAACGGGCTGGCAGCTGGAGTGT	6900
Db	6841	TGAGTACAACTCAGTGGCTGCTCATCAAGCCCTACAAACGGGCTGGCAGCTGGAGTGT	6900
QY	6901	CAGGTACCGCTACGATGGCTGGGGCGGGGTGTCAGCAAGAGCAGCCACACGCCACCA	6960
Db	6901	CAGGTACCGCTACGATGGCTGGGGCGGGGTGTCAGCAAGAGCAGCCACACGCCACCA	6960
QY	6961	CCTGAGTCTTTCTATGAGACCTTGACCAACCCCAAGGTCACCCACCTGTACAACCA	7020
Db	6961	CCTGAGTCTTTCTATGAGACCTTGACCAACCCCAAGGTCACCCACCTGTACAACCA	7020
QY	7021	CTCAGCTCTGAGATCACTCCCTCTACTACAGCTTGCAAGACACCTCTTTTGCATGGA	7080
Db	7021	CTCAGCTCTGAGATCACTCCCTCTACTACAGCTTGCAAGACACCTCTTTTGCATGGA	7080
QY	7081	GCTGAGCAGTGTGATGATTTTACATAGCTTTGTACAAACATCGGACCCCTCTTGTGT	7140
Db	7081	GCTGAGCAGTGTGATGATTTTACATAGCTTTGTACAAACATCGGACCCCTCTTGTGT	7140
QY	7141	CTTTAGTGAACAGTGTGATGATCAAGCAATCTGTACACAGCTTATGGGAGATCTA	7200
Db	7141	CTTTAGTGAACAGTGTGATGATCAAGCAATCTGTACACAGCTTATGGGAGATCTA	7200
QY	7201	CATGATATCAACACCCCACTTTCAGATCATATAGGCTACCATGGTGGCCCTCTATGATCC	7260
Db	7201	CATGATATCAACACCCCACTTTCAGATCATATAGGCTACCATGGTGGCCCTCTATGATCC	7260
QY	7261	ACTCACCAAGCTTGTCCATGAGCGCGGAGATTATGATGTGTGGCCGAGACCTGGAC	7320
Db	7261	ACTCACCAAGCTTGTCCATGAGCGCGGAGATTATGATGTGTGGCCGAGACCTGGAC	7320
QY	7321	TAGCCACAGACAGCTGTGGAAGCCTTAGTAGCAGCAAGCTCATGCCCTTTTAACTCT	7380
Db	7321	TAGCCACAGACAGCTGTGGAAGCCTTAGTAGCAGCAAGCTCATGCCCTTTTAACTCT	7380
QY	7381	CTATATGTTCAAAAACAACACCCCATCAGCAACTCCAGGACATCAAGTCTTCATGAC	7440
Db	7381	CTATATGTTCAAAAACAACACCCCATCAGCAACTCCAGGACATCAAGTCTTCATGAC	7440
QY	7441	AGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAACGTGATCCCTGGTTA	7500
Db	7441	AGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAACGTGATCCCTGGTTA	7500
QY	7501	TCCAAAACAGACATGGATGCCATGGAACCCCTCTACAGCTCATCCACACAGATGAA	7560
Db	7501	TCCAAAACAGACATGGATGCCATGGAACCCCTCTACAGCTCATCCACACAGATGAA	7560
QY	7561	AACGACGAGTGGGACAAACAGCAAGTCTATCCCTCGGGTACAGTGTGAAGTACAGAAGCA	7620
Db	7561	AACGACGAGTGGGACAAACAGCAAGTCTATCCCTCGGGTACAGTGTGAAGTACAGAAGCA	7620
QY	7621	GCTCAAGGCTTTGTACCTTAGAACGGTTTCACACAGCTCTATGGTCCACAAATCACCAG	7680
Db	7621	GCTCAAGGCTTTGTACCTTAGAACGGTTTCACACAGCTCTATGGTCCACAAATCACCAG	7680
QY	7681	CTGCCAGCAGGTCCTCAAGAACCAAGAGTTTCATCCAGCGCTCAGTCTTTGGCAAGGG	7740
Db	7681	CTGCCAGCAGGTCCTCAAGAACCAAGAGTTTCATCCAGCGCTCAGTCTTTGGCAAGGG	7740
QY	7741	GGTCAAGTTTGCCTTGAAGGATGGCGGAGTGACACAGACATCATCAGTGTGCCAATGA	7800
Db	7741	GGTCAAGTTTGCCTTGAAGGATGGCGGAGTGACACAGACATCATCAGTGTGCCAATGA	7800
QY	7801	GGATGGCGCAAGGGTTGCTGCCATCTTGAACCATGCCCCACTACCTAGAGAACCTGCACCT	7860
Db	7801	GGATGGCGCAAGGGTTGCTGCCATCTTGAACCATGCCCCACTACCTAGAGAACCTGCACCT	7860
QY	7861	CACCATTTGATGGGTGGATACCATTTACTTTGTGAAACAGGACCTTTCAGAGGTGACCT	7920
Db	7861	CACCATTTGATGGGTGGATACCATTTACTTTGTGAAACAGGACCTTTCAGAGGTGACCT	7920
QY	7921	GGCATCTCTGGGCTCAGTGGGGGGCGGCAACCTTGGAGAAATGGGGTCAACGCTCACTGT	7980
Db	7921	GGCATCTCTGGGCTCAGTGGGGGGCGGCAACCTTGGAGAAATGGGGTCAACGCTCACTGT	7980
QY	7981	GTCCAGATCAACAACAGTACTTAATGGCAGGACTAGACGCTACACAGATCCAGCTCCA	8040
Db	7981	GTCCAGATCAACAACAGTACTTAATGGCAGGACTAGACGCTACACAGATCCAGCTCCA	8040
QY	8041	GTTACGGGGCACTGTGCTTGAACACACGCTACCGGCAACCTTGGATGAGAGAGCAACG	8100
Db	8041	GTTACGGGGCACTGTGCTTGAACACACGCTACCGGCAACCTTGGATGAGAGAGCAACG	8100
QY	8101	GGTCTCTGGAGCTGGCCCCGCGCAGAGCCGTGGCCAAAGCTGGGCCCGGAGCAGAG	8160
Db	8101	GGTCTCTGGAGCTGGCCCCGCGCAGAGCCGTGGCCAAAGCTGGGCCCGGAGCAGAG	8160
QY	8161	ACTCGGGGAAGGGAGAAAGGCTTCGGGCTGGACAGAGGGGGAGAAACAGAGGTGCT	8220
Db	8161	ACTCGGGGAAGGGAGAAAGGCTTCGGGCTGGACAGAGGGGGAGAAACAGAGGTGCT	8220
QY	8221	GAGCACAGGCGGGTGCACAGGCTTACACGCTTTTTCGTGATCTCTGTGAGCAGTACCC	8280
Db	8221	GAGCACAGGCGGGTGCACAGGCTTACACGCTTTTTCGTGATCTCTGTGAGCAGTACCC	8280
QY	8281	AGAACTGTGACAGCGGCCAAACAACTCCATTTCATGACAGAGGAGATGGCCGGAG	8340
Db	8281	AGAACTGTGACAGCGGCCAAACAACTCCATTTCATGACAGAGGAGATGGCCGGAG	8340
QY	8341	GTGACAGAGAGGAC 8354	
Db	8341	GTGACAGAGAGGAC 8354	

RESULT 2
AX600210
LOCUS
DEFINITION
AX600210
ACCESSION
VERSION
AX600210.1
KEYWORDS
SOURCE
Homo sapiens (human)

AX600210
Sequence 22 from Patent WO02072830.
PAT 14-FEB-2003

ORGANISM	Homo sapiens	Db
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	Qy
AUTHORS	1. Yue, H., Yao, M.G., Ison, C.H., Lu, Y., Warren, B.A., Elliott, V.S., Baughn, M.R., Ding, L., Xu, Y., Gietzen, K.J., Tang, T.Y., Lei, P.G., Duggan, B.M., Burford, N., Lu, D.A., Richardson, T.W., Iran, U.K., Khare, R. and Wallia, N.K.	Db
TITLE	Proteins associated with cell growth, differentiation, and death	Qy
JOURNAL	Patent: WO 02072830-A 22 19-SEP-2002;	Db
FEATURES	Incyte Genomics, Inc. (US)	Qy
source	Location/Qualifiers	Db
	1..8645	Qy
	/organism="Homo sapiens"	Qy
	/mol_type="unassigned DNA"	Qy
	/db_xref="taxon:9606"	Qy
	/note="Incyte ID No: 7498573CB1"	Qy
ORIGIN		Qy
	Query Match 98.5%; Score 8225; DB 6; Length 8645;	Qy
	Best local Similarity 99.4%; Pred. No. 0;	Qy
	Matches 8300; Conservative 0; Mismatches 21; Indels 33; Gaps 3;	Qy
Qy	1 GTTCTGGATGTGAGGAGCGCGGCGGAGCCATGGACGTGAAGGAGAGGAGCCTTA 60	Qy
Db	83 GTTCTGGATGTGAGGAGCGCGGCGGAGCCATGGACGTGAAGGAGAGGAGCCTTA 142	Qy
Qy	61 CGCTCGCTGACCCGGCGCGGAGCGCGGAGCGCGGTACACGAGCTGTCGCGGAGAG 120	Qy
Db	143 CGCTCGCTGACCCGGCGCGGAGCGCGGAGCGCGGTACACGAGCTGTCGCGGAGAG 202	Qy
Qy	121 CGAGGAGGCGAAGCGCGGAGAGATCGTACAGCTCCAGCGAGACCTGAGGCGCTACGA 180	Qy
Db	203 CGAGGAGGCGAAGCGCGGAGAGATCGTACAGCTCCAGCGAGACCTGAGGCGCTACGA 262	Qy
Qy	181 CCAGGACGCCCGCTAGCTATGCGAGCGCGGTCAAGGACATTTGCGCGAGGAGGCCGA 240	Qy
Db	263 CCAGGACGCCCGCTAGCTATGCGAGCGCGGTCAAGGACATTTGCGCGAGGAGGCCGA 322	Qy
Qy	241 GGAATTCGCGGACAGGTGCGCACTTACCCCTCGGAGGTGCGGCTGGAAGATTAAC 300	Qy
Db	323 GGAATTCGCGGACAGGTGCGCACTTACCCCTCGGAGGTGCGGCTGGAAGATTAAC 382	Qy
Qy	301 GCCCCTCAGCGGACCTGTATCCGAGCAGACATTTGGGCTGCCCCAATGCGGCTACTCAT 360	Qy
Db	383 GCCCCTCAGCGGACCTGTATCCGAGCAGACATTTGGGCTGCCCCAATGCGGCTACTCAT 442	Qy
Qy	361 GGGGCTGGCTCTGATGCGGACATGAGGCTGACACGGTGTGTCCTGAGCACCCCGT 420	Qy
Db	443 GGGGCTGGCTCTGATGCGGACATGAGGCTGACACGGTGTGTCCTGAGCACCCCGT 502	Qy
Qy	421 CGCTGTGCGGCGCGGAGCAGCGGTGAGGCGGAGCTCTGCGCTGTCAGAGCGGGCCAA 480	Qy
Db	503 CGCTGTGCGGCGCGGAGCAGCGGTGAGGCGGAGCTCTGCGCTGTCAGAGCGGGCCAA 562	Qy
Qy	481 TTCCAATCTCACTCAACCGACACCGGATGAAAACTGAGACTGATCATCCGGGCGG 540	Qy
Db	563 TTCCAATCTCACTCAACCGACACCGGATGAAAACTGAGACTGATCATCCGGGCGG 616	Qy
Qy	541 CCTCAGAACACCGGGGCTCCGAGCGCGCGCGCGCTCTCGCAGCGCCACACCC 600	Qy
Db	617 CCTCAGAACACCGGGGCTCCGAGCGCGCGCGCGCTCTCGCAGCGCCACACCC 676	Qy
Qy	601 CAACAGCACACCGCGCTCCATTAATCCCTGAAACCGGGGCAACTTTCACGCGGAGGAG 660	Qy
Db	677 CAACAGCACACCGCGCTCCATTAATCCCTGAAACCGGGGCAACTTTCACGCGGAGGAG 736	Qy
Qy	661 CAACCCAGCCCGCGGCTCCGAGCGCGCGCGCGCTCTCGGAGAGCCCGCTCCGCGGCGG 720	Qy
Db	737 CAACCCAGCCCGCGGCTCCGAGCGCGCGCGCGCTCTCGGAGAGCCCGCTCCGCGGCGG 796	Qy
Qy	721 CCAGGAGCTGCCACGCCCGGAGGAACTGGCTCTCAACAGCAAACTCCCTCGGAGAC 780	Qy
Db		Qy
797	CCAGGAGCTGCCACGCCCGGAGGAACTGGCTCTCAACAGCAAACTCCCTCGGAGAC 856	Qy
781	CAGGAACCTAGGCAAGCAGCCATTCCTAGGACATTCAGGAGCAACCTCATTGACATGA 840	Qy
857	CAGGAACCTAGGCAAGCAGCCATTCCTAGGACATTCAGGAGCAACCTCATTGACATGA 916	Qy
841	CATTTCGGCGCTCCCGCATGATGGGCTTACAGTACCGGCACTTCCTCTTCAAGCC 900	Qy
917	CATTTCGGCGCTCCCGCATGATGGGCTTACAGTACCGGCACTTCCTCTTCAAGCC 976	Qy
901	TGAGGCACTCCCGCTCTTCTGCACACATCACAGGTTACCCACTGAGTCCAGCAC 960	Qy
977	TGAGGCACTCCCGCTCTTCTGCACACATCACAGGTTACCCACTGAGTCCAGCAC 1036	Qy
961	AGTGTACTCTCTCCCGCCCGACCCCTGCGCGGAGCACTTCGCGCGGCGGCTTTAA 1020	Qy
1037	AGTGTACTCTCTCCCGCCCGACCCCTGCGCGGAGCACTTCGCGCGGCGGCTTTAA 1096	Qy
1021	CCTCAAGAGCCCTCAAGTACTGTAACTGGAGTGGCAGCCCTGAGGCGCATGTCAT 1080	Qy
1097	CCTCAAGAGCCCTCAAGTACTGTAACTGGAGTGGCAGCCCTGAGGCGCATGTCAT 1156	Qy
1081	CTCAGCACTCTGGTCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140	Qy
1157	CTCAGCACTCTGGTCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1210	Qy
1141	CTGGAACCTGAGCGAGTGGAGGCGAGATGTATGAGATCAAGGAGCACAGCCAGCAG 1200	Qy
1211	CTGGAACCTGAGCGAGTGGAGGCGAGATGTATGAGATCAAGGAGCACAGCCAGCAG 1270	Qy
1201	TTGGCTGTGCGCAACCGAGCTCTCCCTATACCCCTCAGGCGGCACTGCTTACAGACCC 1260	Qy
1271	TTGGCTGTGCGCAACCGAGCTCTCCCTATACCCCTCAGGCGGCACTGCTTACAGACCC 1330	Qy
1261	TCAGAGGAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1320	Qy
1331	TCAGAGGAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1390	Qy
1321	TTTCATAGATTCCTGAGAAATGATGTTGGGAGGCGAGCTCCAGAGATTCCTCTGG 1380	Qy
1391	TTTCATAGATTCCTGAGAAATGATGTTGGGAGGCGAGCTCCAGAGATTCCTCTGG 1450	Qy
1381	CACCTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGATCTGAAATTCATATGTC 1440	Qy
1451	CACCTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGATCTGAAATTCATATGTC 1510	Qy
1441	TCGGGAAAGGAGCCCTGGTTGGCATTTATGGCAGAAAGGCTCCCTCTTCAATAC 1500	Qy
1511	TCGGGAAAGGAGCCCTGGTTGGCATTTATGGCAGAAAGGCTCCCTCTTCAATAC 1570	Qy
1501	ACAGTTTGAATTTGGAGCTGCTGGATGGCAGAGGCTCCCTAACCCAGGAGGCGGAG 1560	Qy
1571	ACAGTTTGAATTTGGAGCTGCTGGATGGCAGAGGCTCCCTAACCCAGGAGGCGGAG 1630	Qy
1561	CCTAGAGGAGCCCGCGCCAGTCTCGGGAACTGTGCCCCCTCCAGCCATGAGACAGG 1620	Qy
1631	CCTAGAGGAGCCCGCGCCAGTCTCGGGAACTGTGCCCCCTCCAGCCATGAGACAGG 1690	Qy
1621	CTTCATCCAGATTTGGATTCAGGAATCTGCACTTGGCTTTTACATGACGGAAGGA 1680	Qy
1691	CTTCATCCAGATTTGGATTCAGGAATCTGCACTTGGCTTTTACATGACGGAAGGA 1750	Qy
1681	GTCAAGTGGTTTCTCTTCTCACACTGCCATTAAGTTCGGTGGATTAATGCCCCAGCAA 1740	Qy
1751	GTCAAGTGGTTTCTCTTCTCACACTGCCATTAAGTTCGGTGGATTAATGCCCCAGCAA 1810	Qy
1741	CTGCTATGGCAATGGTGAATCTGGAACCTGCACTGCTTCTGCTGGTTCCTGGG 1800	Qy
1811	CTGCTATGGCAATGGTGAATCTGGAACCTGCACTGCTTCTGCTGGTTCCTGGG 1870	Qy
1801	CCCCAGCTGTGGCAGAGCCTCTGCCCCGCTCTGTAGCGGAAATGCCCATAATGAA 1860	Qy
1871	CCCCAGCTGTGGCAGAGCCTCTGCCCCGCTCTGTAGCGGAAATGCCCATAATGAA 1930	Qy

Qy	1861	AGCAGATGCTTGTGCCACAGTGGCTGGAAAGCGGTGAGTGGATGTGCCCAACCA	1920
Db	1931	AGGCAGATGCTTGTGCCACAGTGGCTGGAAAGCGGTGAGTGGATGTGCCCAACCA	1990
Qy	1921	GTGTATCGATGTGGCTGTCAGCAACCATGCACTGCATCAAGGGACCTGCAATGCAA	1980
Db	1991	GTGTATCGATGTGGCTGTCAGCAACCATGCACTGCATCAAGGGACCTGCAATGCAA	2050
Qy	1981	CCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGTCATGGACCCCAATGTTCAAG	2040
Db	2051	CCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGTCATGGACCCCAATGTTCAAG	2110
Qy	2041	CCGGGTGTCTGGTGTAGAGCGCAATGCCATTCCTTGTGGGATGGGGGACCAACTG	2100
Db	2111	CCGGGTGTCTGGTGTAGAGCGCAATGCCATTCCTTGTGGGATGGGGGACCAACTG	2170
Qy	2101	CGAGACCCCGAGGGCCACATGCTTTAGACCAGTGTTCAGGGCCAGGGAACCTTCCTCCCGGA	2160
Db	2171	CGAGACCCCGAGGGCCACATGCTTTAGACCAGTGTTCAGGGCCAGGGAACCTTCCTCCCGGA	2230
Qy	2161	CACGGGCTTTGAGCTGTGACCCAGCTGGAATGGACACGACTGTTCATCAGATCTG	2220
Db	2231	CACGGGCTTTGAGCTGTGACCCAGCTGGAATGGACACGACTGTTCATCAGATCTG	2290
Qy	2221	TGCTGCCAGCTGTGTGGTGGCCATGCGCTGTGCTAGGGGGACACTGCCGCTGCCAGATGG	2280
Db	2291	TGCTGCCAGCTGTGTGGTGGCCATGCGCTGTGCTAGGGGGACACTGCCGCTGCCAGATGG	2350
Qy	2281	CTGGATGGGGGACGCTTGCACACAGCGGGCTTGCACCCGGCTGTGCCAGCATGGAC	2340
Db	2351	CTGGATGGGGGACGCTTGCACACAGCGGGCTTGCACCCGGCTGTGCCAGCATGGAC	2410
Qy	2341	CTCCGGCGAGGCAAGTCCGAGTGCACCCCTGCTGGAATGGGCAACATGCACCACTCGC	2400
Db	2411	CTCCGGCGAGGCAAGTCCGAGTGCACCCCTGCTGGAATGGGCAACATGCACCACTCGC	2470
Qy	2401	TCACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTGTGCAATGGCAACGGCAG	2460
Db	2471	TCACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTGTGCAATGGCAACGGCAG	2530
Qy	2461	ATGTACCTTAGACCTGTAATGGTTGGCACTGGCTTGCACGCTGGGCTGGAGAGGAGCTGG	2520
Db	2531	ATGTACCTTAGACCTGTAATGGTTGGCACTGGCTTGCACGCTGGGCTGGAGAGGAGCTGG	2590
Qy	2521	CTGTGACACTTCCATGTGAGACTGCTCGGCTGACAGCAAGAACAATGATGGAGATGGCCT	2580
Db	2591	CTGTGACACTTCCATGTGAGACTGCTCGGCTGACAGCAAGAACAATGATGGAGATGGCCT	2650
Qy	2581	GGTGGACTGATGGAACCTGACTGCTGCTCCAGCCCTGTGCCATATCAACCCGCTGTG	2640
Db	2651	GGTGGACTGATGGAACCTGACTGCTGCTCCAGCCCTGTGCCATATCAACCCGCTGTG	2710
Qy	2641	CCATTGGCTCCCTAAACCTCTGGACATATCCAGGAGACACAGTCCCTGTGTGCACAGCA	2700
Db	2711	CCATTGGCTCCCTAAACCTCTGGACATATCCAGGAGACACAGTCCCTGTGTGCACAGCA	2770
Qy	2701	GAACCTACACTCCTTCTATGACCCGATCAAGTTCCCTGTGGGACGGGACAGCACGACAT	2760
Db	2771	GAACCTACACTCCTTCTATGACCCGATCAAGTTCCCTGTGGGACGGGACAGCACGACAT	2830
Qy	2761	AATCCCGGGGAGAACCCCTTTGATGAGGGCATGCTTGTGTATTTCGTGGCCAGTGAT	2820
Db	2831	AATCCCGGGGAGAACCCCTTTGATGAGGGCATGCTTGTGTATTTCGTGGCCAGTGAT	2890
Qy	2821	GACATCAGATGGAAACCCCTCTGGTTGGTGTGAAATCAGTTTTTGTCAATAACCCCTCTCT	2880
Db	2891	GACATCAGATGGAAACCCCTCTGGTTGGTGTGAAATCAGTTTTTGTCAATAACCCCTCTCT	2950
Qy	2881	TGGATATACAAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGTACAAATGGCGCATCTC	2940
Db	2951	TGGATATACAAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGTACAAATGGCGCATCTC	3010

[illegible]

Db 4070 CCCCCTTTGATGACATCTCGCTGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACCAA 4129
 Qy 4081 TCCCGAGGGTATTACAGTGGCAAGCTTTGGGCTGATCTACTTCGTGGATGGCCACCATGAT 4140
 Db 4130 TCCCGAGGGCATTTACAGTGGCAAGCTTTGGGCTGATCTACTTCGTGGATGGCCACCATGAT 4189
 Qy 4141 CAGAGGCATCGATCAGAAATGGGATCATCTCCACCCCTGCTCGGCTCTTAATGATCTCACATC 4200
 Db 4190 CAGAGGCATCGATCAGAAATGGGATCATCTCCACCCCTGCTCGGCTCTTAATGATCTCACATC 4249
 Qy 4201 AGCCCGGCCACTCAGCTGATCTGTGATGATATTTCCAGGTGAAGATCGAGTGGCC 4260
 Db 4250 AGCCCGGCCACTCAGCTGATCTGTGATGATATTTCCAGGTTCACCTGGAGTGGCC 4309
 Qy 4261 CACAGACTTAGCCATCAACCCCAATGGCAAACTCACTTTATGCTTCGACAAACAATGTGGT 4320
 Db 4310 CACAGACTTAGCCATCAACCCCAATGGCAAACTCACTTTATGCTTCGACAAACAATGTGGT 4369
 Qy 4321 CCGTCAAAATCTCTGAAAACCCAGGTGGCATTTGTCCCGGGAGGCCATGCACTGCCA 4380
 Db 4370 CCGTCAAAATCTCTGAAAACCCAGGTGGCATTTGTCCCGGGAGGCCATGCACTGCCA 4429
 Qy 4381 GGTCCCTGGCATTTGACCACTTCTCTGTAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTC 4440
 Db 4430 GGTCCCTGGCATTTGACCACTTCTCTGTAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTC 4489
 Qy 4441 AGCCACCCGCTTTGGCTTTTACAAATGGGGTCTGTATATTCGTGAGACTGATGAAA 4500
 Db 4490 AGCCACCCGCTTTGGCTTTTACAAATGGGGTCTGTATATTCGTGAGACTGATGAAA 4549
 Qy 4501 AAGATCAACCGCATCAGCGAGTCAACCACTAGTGGAGAGATCTCACTGCTGCTGGGC 4560
 Db 4550 AAGATCAACCGCATCAGCGAGTCAACCACTAGTGGAGAGATCTCACTGCTGCTGGGC 4609
 Qy 4561 CCCCAGTGGCTGTGATCTGATAAAATGATGCCAACTGTGATTTTCTGGAGACGATGG 4620
 Db 4610 CCCCAGTGGCTGTGATCTGATAAAATGATGCCAACTGTGATTTTCTGGAGACGATGG 4669
 Qy 4621 TTATGCCAAGATGCAAAATGATAACCCCATCTTCCTTGGCTGTGCTGTGATGGGA 4680
 Db 4670 TTATGCCAAGATGCAAAATGATAACCCCATCTTCCTTGGCTGTGCTGTGATGGGA 4729
 Qy 4681 GCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTT 4740
 Db 4730 GCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTT 4789
 Qy 4741 CCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACACGAGCTCTATCTGTT 4800
 Db 4790 CCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACACGAGCTCTATCTGTT 4849
 Qy 4801 TGATACACCCGCGAGACCTGTACACCAAGCCTGCCACAGGACTACCTGTACAA 4860
 Db 4850 TGATACACCCGCGAGACCTGTACACCAAGCCTGCCACAGGACTACCTGTGTACAA 4909
 Qy 4861 CTTCACTACACTGGGGACGGCGACATCACAATCTCATCACAGAACAAATGCAACATGGT 4920
 Db 4910 CTTCACTACACTGGGGACGGCGACATCACAATCTCATCACAGAACAAATGCAACATGGT 4969
 Qy 4921 AAATGTCGCGAGACTCTACTGGGATGCCCTCTGCTGGTGTGCTGAGATGGCCAGGT 4980
 Db 4970 AAATGTCGCGAGACTCTACTGGGATGCCCTCTGCTGGTGTGCTGAGATGGCCAGGT 5029
 Qy 4981 GTACTGGGTGACCATGGGCACCAACAGTGCATCAAGATGTGACCAACAAGGACACGA 5040
 Db 5030 GTACTGGGTGACCATGGGCACCAACAGTGCATCAAGATGTGACCAACAAGGACACGA 5089
 Qy 5041 GTTGCCCATGATGACATACCAATGGCAATTCGGGCTTCTGGCCATCCAAAAGCAATGAAA 5100
 Db 5090 GTTGCCCATGATGACATACCAATGGCAATTCGGGCTTCTGGCCATCCAAAAGCAATGAAA 5149
 Qy 5101 CGGATGGACAAATTTATGATGACGAGCTTTGGCCGCTGTGACAAATGTGACCTTCCC 5160

Db 5150 CGGATGGACAAACATTTTATGATGACACAGCTTTTGGCCGCTTGACAAATGTGACCTTCCC 5209
 Qy 5161 TACTGCCAGGTGAGCAGTTTCCGAAGTATACAGACAGTTTCAGTGCAATGTCCAGGTAGA 5220
 Db 5210 TACTGCCAGGTGAGCAGTTTCCGAAGTATACAGACAGTTTCAGTGCAATGTCCAGGTAGA 5269
 Qy 5221 GACCTCCAGCAAGGATGATTCACCATAAACCAACCACTGTCTGCTCAGGCGCCTTTCTA 5280
 Db 5270 GACCTCCAGCAAGGATGATTCACCATAAACCAACCACTGTCTGCTCAGGCGCCTTTCTA 5329
 Qy 5281 CACACTGCTCAGACCAAGTCCGGAACAGCTACTACATCGGGCGGATGGCTCTTGGC 5340
 Db 5330 CACACTGCTCAGACCAAGTCCGGAACAGCTACTACATCGGGCGGATGGCTCTTGGC 5389
 Qy 5341 GGTGCTGCTGGGCAACCGCATGGAGTGGCGCTGAGAGCTGAGCCGCCACTTGTCTGGCTGG 5400
 Db 5390 GGTGCTGCTGGGCAACCGCATGGAGTGGCGCTGAGAGCTGAGCCGCCACTTGTCTGGCTGG 5449
 Qy 5401 CACCTGCAACCCCAACCGTGGGCAAGGAAATGTACGCTGCCCATCGAACCGGCTCAA 5460
 Db 5450 CACCTGCAACCCCAACCGTGGGCAAGGAAATGTACGCTGCCCATCGAACCGGCTCAA 5509
 Qy 5461 CTTGTTGGAGTGGCGCAGCGCAAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGCGG 5520
 Db 5510 CTTGTTGGAGTGGCGCAGCGCAAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGCGG 5569
 Qy 5521 CCGGCTGCGGGTGCAAAACCGAAATCTCTATCTCTGGAATTTGATCGCGTAACACGCA 5580
 Db 5570 CCGGCTGCGGGTGCAAAACCGAAATCTCTATCTCTGGAATTTGATCGCGTAACACGCA 5629
 Qy 5581 AGAGAGATCTATGATGACCCAGCAAGTTCACCTTCGGATTCGTGACGACATCTCCCC 5640
 Db 5630 AGAGAGATCTATGATGACCCAGCAAGTTCACCTTCGGATTCGTGACGACATCTCCCC 5689
 Qy 5641 GCGGCCACCCCTCTGGTCAACCCAGCAGGCTGAAATGGTGTCAACGTGACATCTCCCC 5700
 Db 5690 GCGGCCACCCCTCTGGTCAACCCAGCAGGCTGAAATGGTGTCAACGTGACATCTCCCC 5749
 Qy 5701 TGGGGTTACATGCTGGCATCCAGGGGATCATGTCTGAAAGAAATGGAATACGACCA 5760
 Db 5750 TGGGGTTACATGCTGGCATCCAGGGGATCATGTCTGAAAGAAATGGAATACGACCA 5809
 Qy 5761 GCGGGCCGCATCACATCCAGGATCTTGGCTGATGGGAAGACATGGAGCTACATATCTT 5820
 Db 5810 GCGGGCCGCATCACATCCAGGATCTTGGCTGATGGGAAGACATGGAGCTACATATCTT 5869
 Qy 5821 AGAGAGTCCATGCTGCTGCTACTACAGCCAGGAGGATATATCTTTGATGTTGCAAA 5880
 Db 5870 AGAGAGTCCATGCTGCTGCTACTACAGCCAGGAGGATATATCTTTGATGTTGCAAA 5929
 Qy 5881 GAATGACCGCTCTCTTGTGACGATGCCAACTGGCGGCGACACATAGAGACCAT 5940
 Db 5930 GAATGACCGCTCTCTTGTGACGATGCCAACTGGCGGCGACACATAGAGACCAT 5989
 Qy 5941 CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCTCCTGAGGCAATGCTCAGTCA 6000
 Db 5990 CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCTCCTGAGGCAATGCTCAGTCA 6049
 Qy 6001 ACAGGACTTCACTGAGGATGGGACCTCTTACACCTTCTACCTGGGACATGGCCGAG 6060
 Db 6050 ACAGGACTTCACTGAGGATGGGACCTCTTACACCTTCTACCTGGGACATGGCCGAG 6109
 Qy 6061 GGTGATATACAGATGGCAAACTGTCAAAGTGGCAGAGCGCTCTTATGACACCAACAA 6120
 Db 6110 GGTGATATACAGATGGCAAACTGTCAAAGTGGCAGAGCGCTCTTATGACACCAACAA 6169
 Qy 6121 GGTGAGTTTCACTATGACGAGACGGCAGGATGTCTGAAGACCATCAACCTACAGATGA 6180
 Db 6170 GGTGAGTTTCACTATGACGAGACGGCAGGATGTCTGAAGACCATCAACCTACAGATGA 6229
 Qy 6181 GGGCTTCACTGACCAATCCGCTACCGTACGATTGGGCCCCCTGATTGACCCGACAGATCTT 6240
 Db 6230 GGGCTTCACTGACCAATCCGCTACCGTACGATTGGGCCCCCTGATTGACCCGACAGATCTT 6289

QY 6241 CCGCTTCACTGAGGAGGATGCTCAACGCCCGTTTTGACTACAACTATGACAAACAGCTT 6300
 Db |||||
 QY 6290 CCGCTTCACTGAGGAGGATGCTCAACGCCCGTTTTGACTACAACTATGACAAACAGCTT 6349
 Db |||||
 QY 6301 CCGGTGACCAAGCATGACGAGTGTGATCAACGAGACCCCACTGCCATTTGATCTCTATCG 6360
 Db |||||
 QY 6350 CCGGTGACCAAGCATGACGAGTGTGATCAACGAGACCCCACTGCCATTTGATCTCTATCG 6409
 Db |||||
 QY 6361 CTATGATGATGTGACGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCAATTACTATGA 6420
 Db |||||
 QY 6410 CTATGATGATGTGACGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCAATTACTATGA 6469
 Db |||||
 QY 6421 CATTAACAGATCATCAACAGAGTGTGATCAACGAGACCCCACTGCCATTTGATGATATGG 6480
 Db |||||
 QY 6470 CATTAACAGATCATCAACAGAGTGTGATCAACGAGACCCCACTGCCATTTGATGATATGG 6529
 Db |||||
 QY 6481 CAGGATGAAGGAAGTGCAGTATGAGATCTTCGGCTCGCTCATGTACTGGATGACCGTCCA 6540
 Db |||||
 QY 6530 CAGGATGAAGGAAGTGCAGTATGAGATCTTCGGCTCGCTCATGTACTGGATGACCGTCCA 6589
 Db |||||
 QY 6541 GTATGATTAACATGGGGCGAGTAGTGAAGAAGAGCTGAAGTGAAGACCTACGCCAATAC 6600
 Db |||||
 QY 6590 GTATGATTAACATGGGGCGAGTAGTGAAGAAGAGCTGAAGTGAAGACCTACGCCAATAC 6649
 Db |||||
 QY 6601 CACTCGCTACTCTCTATGATGATGATCTGACGGCCAGCTGCAGACAGTCTCCATCAATGA 6660
 Db |||||
 QY 6650 CACTCGCTACTCTCTATGATGATGATCTGACGGCCAGCTGCAGACAGTCTCCATCAATGA 6709
 Db |||||
 QY 6661 CAAGCCACTCTGGGCTACAGTACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGG 6720
 Db |||||
 QY 6710 CAAGCCACTCTGGGCTACAGTACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGG 6769
 Db |||||
 QY 6721 GAACAGTGCAGGGCTCAACCACTACGGTATGATCGCGACCGCATCACTCGGCTGGG 6780
 Db |||||
 QY 6770 GAACAGTGCAGGGCTCAACCACTACGGTATGATCGCGACCGCATCACTCGGCTGGG 6829
 Db |||||
 QY 6781 TGACGTGCAATACAGATGAGTGGTCTTCTGAGGAGCGGGCGGTGATATCTT 6840
 Db |||||
 QY 6830 TGACGTGCAATACAGATGAGTGGTCTTCTGAGGAGCGGGCGGTGATATCTT 6889
 Db |||||
 QY 6841 TGAGTACAACTCAGTGGCTGCTCATCAAGCCCTACAAACCGGCTGACGCTGGAGTGT 6900
 Db |||||
 QY 6890 TGAGTACAACTCAGTGGCTGCTCATCAAGCCCTACAAACCGGCTGACGCTGGAGTGT 6949
 Db |||||
 QY 6901 CAGGTACCGCTACGATGCGCTGGGGCGGCTGTCCAGCAAGAGCAGCCACAGCCACCA 6960
 Db |||||
 QY 6950 CAGGTACCGCTACGATGCGCTGGGGCGGCTGTCCAGCAAGAGCAGCCACAGCCACCA 7009
 Db |||||
 QY 6961 CCGTCAAGTCTTCTATGAGACCTGACCAACCCCAAGGTCAACCACTGTACACCA 7020
 Db |||||
 QY 7010 CCGTCAAGTCTTCTATGAGACCTGACCAACCCCAAGGTCAACCACTGTACACCA 7069
 Db |||||
 QY 7021 CTCAGACTGAGATCACTCCCTCTATAGACTTTCAGGACAGACCTTTTGGCATGGA 7080
 Db |||||
 QY 7070 CTCAGACTGAGATCACTCCCTCTATAGACTTTCAGGACAGACCTTTTGGCATGGA 7129
 Db |||||
 QY 7081 GCTGAGCAGTGGTGTGATGATTTTACATAGCTTTGACAACTACGGAACCCCTCTTGTGT 7140
 Db |||||
 QY 7130 GCTGAGCAGTGGTGTGATGATTTTACATAGCTTTGACAACTACGGAACCCCTCTTGTGT 7189
 Db |||||
 QY 7141 CTTTGTGGAACAGTGTGATGATCAAGCAATCTGTACAGAGCTTATGGGAGTCTA 7200
 Db |||||
 QY 7190 CTTTGTGGAACAGTGTGATGATCAAGCAATCTGTACAGAGCTTATGGGAGTCTA 7249
 Db |||||
 QY 7201 CATGATACCAACCCCACTTTCAGATCATATAGGCTTACCATGGTGGCTCTATGATCC 7260
 Db |||||
 QY 7250 CATGATACCAACCCCACTTTCAGATCATATAGGCTTACCATGGTGGCTCTATGATCC 7309
 Db |||||
 QY 7261 ACTCAACAGCTTGTCCATGCGGCGGAGATTTATGATGTGTGCGCGGAGCGGTGGAC 7320
 Db |||||
 QY 7310 ACTCAACAGCTTGTCCATGCGGCGGAGATTTATGATGTGTGCGGCGGAGCGGTGGAC 7369
 Db |||||

QY 7321 TAGCCGAGACCAAGAGCTGTGGAAGCACCTTAGTAGCAGCAACGCTCATGCCCTTTAACTCT 7380
 Db |||||
 QY 7370 TAGCCGAGACCAAGAGCTGTGGAAGCACCTTAGTAGCAGCAACGCTCATGCCCTTTAACTCT 7429
 Db |||||
 QY 7381 CTATATGTTTCAAAAAACAACCCCATCAGCAACTCCCAGGACATCAAGTGTCTCATGAC 7440
 Db |||||
 QY 7430 CTATATGTTTCAAAAAACAACCCCATCAGCAACTCCCAGGACATCAAGTGTCTCATGAC 7489
 Db |||||
 QY 7441 AGATGTTTACAGCTGGTCTCTCACCTTTGGATTCCAGCTACACAAGCTGATCCCTGGTTA 7500
 Db |||||
 QY 7490 AGATGTTTACAGCTGGTCTCTCACCTTTGGATTCCAGCTACACAAGCTGATCCCTGGTTA 7549
 Db |||||
 QY 7501 TCCCAAAACCAAGCATGGATGCCATCGAACCCCTCCTACGAGCTCATCCACACAGATGAA 7560
 Db |||||
 QY 7550 TCCCAAAACCAAGCATGGATGCCATCGAACCCCTCCTACGAGCTCATCCACACAGATGAA 7609
 Db |||||
 QY 7561 AACGAGGAGTGGGACAAACAGCAAGTCTATCTCGGGGTACAGTGTGAAGTACAGAGCA 7620
 Db |||||
 QY 7610 AACGAGGAGTGGGACAAACAGCAAGTCTATCTCGGGGTACAGTGTGAAGTACAGAGCA 7669
 Db |||||
 QY 7621 GCTCAAGGCTTTGTCAACCTTAGAACGGTTTGACAGCTCTATGGCTCCACAATCACCAAG 7680
 Db |||||
 QY 7670 GCTCAAGGCTTTGTCAACCTTAGAACGGTTTGACAGCTCTATGGCTCCACAATCACCAAG 7729
 Db |||||
 QY 7681 CTGCCAGAGGCTCCAAAGACCAAGAAAGTTTGATCCAGCGCTCAGTCTTTGGCAAGGG 7740
 Db |||||
 QY 7730 CTGCCAGAGGCTCCAAAGACCAAGAAAGTTTGATCCAGCGCTCAGTCTTTGGCAAGGG 7789
 Db |||||
 QY 7741 GGTCAAGTGTTCCTTGAAGATGGCGAGTGGCCAGACCAAGACATCATCAGTGTGCCAATCA 7800
 Db |||||
 QY 7790 GGTCAAGTGTTCCTTGAAGATGGCGAGTGGCCAGACCAAGACATCATCAGTGTGCCAATCA 7849
 Db |||||
 QY 7801 GGATGGCGCAAGGGTTGCTGCCATCTTGAACCATGCCCCACTACCTAGAGAAACCTGCACATT 7860
 Db |||||
 QY 7850 GGATGGCGCAAGGGTTGCTGCCATCTTGAACCATGCCCCACTACCTAGAGAAACCTGCACATT 7909
 Db |||||
 QY 7861 CACCATTTGATGGGGTGGATACCCATTTACTTTGTGAAACAGAGACCTTTCAGAGGTGACCT 7920
 Db |||||
 QY 7910 CACCATTTGATGGGGTGGATACCCATTTACTTTGTGAAACAGAGACCTTTCAGAGGTGACCT 7969
 Db |||||
 QY 7921 GGCCATCTCGGCTCAGTGGGGGGCGGCAACCCCTGGAGAAATGGGGTCAACCTCACTGT 7980
 Db |||||
 QY 7970 GGCCATCTCGGCTCAGTGGGGGGCGGCAACCCCTGGAGAAATGGGGTCAACCTCACTGT 8029
 Db |||||
 QY 7981 GTCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGATCCAGTCTCCA 8040
 Db |||||
 QY 8030 GTCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGATCCAGTCTCCA 8089
 Db |||||
 QY 8041 GTACGGGCACTGTGCTTGAACACACGCTACGGGCAACGTTGGATGAGGAGAGGCACG 8100
 Db |||||
 QY 8090 GTACGGGCACTGTGCTTGAACACACGCTACGGGCAACGTTGGATGAGGAGAGGCACG 8149
 Db |||||
 QY 8101 GGTCTCTGGAGCTGGCCCGGACAGAGCCGTGGCCCAAGGCTGGGCCCGGAGCAGAGAG 8160
 Db |||||
 QY 8150 GGTCTCTGGAGCTGGCCCGGACAGAGCCGTGGCCCAAGCGTGGGCCCGGAGCAGAGAG 8209
 Db |||||
 QY 8161 ACTCGGGGAAGGGAGGAGGCTGGGGCTGGGACAGAGGGGGAGGAGAGGAGTGTCT 8220
 Db |||||
 QY 8210 ACTCGGGGAAGGGAGGAGGCTGGGGCTGGGACAGAGGGGGAGGAGAGGAGTGTCT 8269
 Db |||||
 QY 8221 GAGCAGAGGGGGTGCAGGCTACAGCGCTTTTTCGTGATCTCTGTGAGCAGTACCC 8280
 Db |||||
 QY 8270 GAGCAGAGGGGGTGCAGGCTACAGCGCTTTTTCGTGATCTCTGTGAGCAGTACCC 8329
 Db |||||
 QY 8281 AGAACTGTACAGAGCGCCAAACAACTCCACTTTCATGACAGAGGAGATGGCGCGGAG 8340
 Db |||||
 QY 8330 AGAACTGTACAGAGCGCCAAACAACTCCACTTTCATGACAGAGGAGATGGCGCGGAG 8389
 Db |||||
 QY 8341 GTCAACAGAGGAC 8354
 Db |||||
 QY 8390 GTGACAGAGAGGAC 8403
 Db |||||

3832 TTTTACAGACATAGTACAGTCCAGCACACAAATACTACTCTGCGCACAGACCCCATGAGTGG 3891
3904 GGGCGCTCTTCTTTCTGACAGCAACAGCGCGGGTCTTTAAATCAAGTCCACTGTGGT 3963
3992 GGGCGCTCTTCTTTCTGACAGCAACAGCGCGGGTCTTTAAATCAAGTCCACTGTGGT 3951
3964 GGTGAAGACCTTGTCAAGAACTCTGAGGTGGTTCGGGGGACAGGTACCAAGTGCCTCCC 4023
3952 GGTGAAGACCTTGTCAAGAACTCTGAGGTGGTTCGGGGGACAGGTACCAAGTGCCTCCC 4011
4024 CTTTATGACACTCGCTCGGGGATGGTGGGAAGCCACAGAGCCACACTCAACAATCC 4083
4012 CTTTATGACACTCGCTCGGGGATGGTGGGAAGCCACAGAGCCACACTCAACAATCC 4071
4084 CAGGGGT-----ATTACAGTGGCAAGTTTGGGCTGATCTACTTCGTGGATGGCAC 4134
4072 CAGGGGTCCCCCAGGCATTACAGTGGACAGTTTGGGCTGATCTACTTCGTGGATGGCAC 4131
4135 CATGATCAGACGATCGATCGAATGGGATCATCTCCACCTGCTCGGCTCTAAATGATCT 4194
4132 CATGATCAGACGATCGATCGAATGGGATCATCTCCACCTGCTCGGCTCTAAATGATCT 4191
4195 CACATCAGCGCGGCCACTCAGCTGTGATCTGTGATGATATTTCCAGGTAAAGA----- 4249
4192 CACATCAGCGCGGCCACTCAGCTGTGATCTGTGATGATATTTCCAGGTAAAGACAGGT 4251
4250 -----CTGGAGTGGCCACAGACTTAGCCATCAACCCCAATGACAACTTTATGTCCT 4305
4252 TCACCTGAGTGGCCACAGACTTAGCCATCAACCCCAATGACAACTTTATGTCCT 4311
4306 CGACAAACATGTGTCCTGCAAACTCTGAAATCCAGAGTGGGATTTGTCGCGGGAG 4365
4312 CGACAAACATGTGTCCTGCAAACTCTGAAATCCAGAGTGGGATTTGTCGCGGGAG 4371
4366 GCCCATGCACTGCCAGGTCCCTGGCATTTGACCACTTCTGCTAAGCAAGTGGCCATCCA 4425
4372 GCCCATGCACTGCCAGGTCCCTGGCATTTGACCACTTCTGCTAAGCAAGTGGCCATCCA 4431
4426 GCACACCTGTGAGTGCAGCCACCGTCTGGCTGTTTACAAATGGGTTCCTGTATATGC 4485
4432 GCACACCTGTGAGTGCAGCCACCGTCTGGCTGTTTACAAATGGGTTCCTGTATATGC 4491
4486 TGAGACTCATGAGAAAGATCAACCGCATCAGCAGTCAACCACTAGTGGAGATCTC 4545
4492 TGAGACTCATGAGAAAGATCAACCGCATCAGCAGTCAACCACTAGTGGAGATCTC 4551
4546 ACTCGTCTGGGCCCCCAGTGGCTGTGACTGTGATAAAATGATGCCAACTGTGATTGTT 4505
4552 ACTCGTCTGGGCCCCCAGTGGCTGTGACTGTGATAAAATGATGCCAACTGTGATTGTT 4511
4606 TTCTGGAGACATGGTTATGCAAGGATGCAAGTAAATACCCCATCTTCTTGGCTGT 4665
4612 TTCTGGAGACATGGTTATGCAAGGATGCAAGTAAATACCCCATCTTCTTGGCTGT 4671
4666 GTGTGCTGATGGGAGCTCTACGTGGCGGACCTTGGGAAATCCGAATTCGGTTATCCG 4725
4672 GTGTGCTGATGGGAGCTCTACGTGGCGGACCTTGGGAAATCCGAATTCGGTTATCCG 4731
4726 GAAGAACAAAGCTTCTCAACACCCAGAACATGTATGAGTGTCTTCAACAATGACCA 4785
4732 GAAGAACAAAGCTTCTCAACACCCAGAACATGTATGAGTGTCTTCAACAATGACCA 4791
4786 GGAGCTCTATCTGTTGATACACCGCAAGCACTGTACACCCAAAGCCTGCCACAGG 4845
4792 GGAGCTCTATCTGTTGATACACCGCAAGCACTGTACACCCAAAGCCTGCCACAGG 4851
4846 AGACTCTGTACAACTTCACTTACATGCGGAGCGGACATCACTCATACAGACAA 4905
4852 AGACTCTGTACAACTTCACTTACATGCGGAGCGGACATCACTCATACAGACAA 4911
4906 CAATGGCAACATGTAATGTCGCGGAGACTTACTGGGATGCCCTCTGGCTGGTGGT 4965
4912 CAATGGCAACATGTAATGTCGCGGAGACTTACTGGGATGCCCTCTGGCTGGTGGT 4971

4966 CCCAGATGCCAGGTCTACTGGGTGACCATGGGACCAACAGTGCCTCAAGAGTGTGC 5025
4972 CCAGATGCCAGGTGTACTGGGTGACCATGGGACCAACAGTGCCTCAAGAGTGTGC 5031
5026 CACAAAGGACACGAGTTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAAC 5085
5032 CACAAAGGACACGAGTTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAAC 5091
5086 CAAAAGCAATGAACGGATGGAACAACATTTTATGATACGACAGCTTTGGCCGCTGC 5145
5092 CAAAAGCAATGAACGGATGGAACAACATTTTATGATACGACAGCTTTGGCCGCTGC 5151
5146 AAATGTGACTTCCCTTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCACT 5205
5152 AAATGTGACTTCCCTTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCACT 5211
5206 GATGTCCAGGTAGAGACTCCAGCAAGATGATCAACATAACCAACACTTGTCTGC 5265
5212 GATGTCCAGGTAGAGACTCCAGCAAGATGATCAACATAACCAACACTTGTCTGC 5271
5266 CTCAGCGCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACTCGGGC 5325
5272 CTCAGCGCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACTCGGGC 5331
5326 CGATGGCTCTTTCGGCTGCTGCTGCGCAACGGCATGGAGTGGCGCTGCAGACTGAGCC 5385
5332 CGATGGCTCTTTCGGCTGCTGCTGCGCAACGGCATGGAGTGGCGCTGCAGACTGAGCC 5391
5386 CCACCTGTGCTGCTGCAACCGCTCAACCGCTGCGGCAAGAGGATGTCACTGCTGCCAT 5445
5392 CCACCTGTGCTGCTGCAACCGCTCAACCGCTGCGGCAAGAGGATGTCACTGCTGCCAT 5451
5446 CGACAAAGGCTTCAACCTGGTGGTGGCGCGCAGGCAAGAGAGGCTCGGGGCGAGGT 5505
5452 CGACAAAGGCTTCAACCTGGTGGTGGCGCGCAGGCAAGAGAGGCTCGGGGCGAGGT 5511
5506 CACTGTCTTGTGGCGCGCGCTGCGGGTG-----CACAAACCGAAATCTCTATCTCT 5556
5512 CACTGTCTTGTGGCGCGCGCTGCGGGTG-----CACAAACCGAAATCTCTATCTCT 5571
5557 GCACCTTTGATCCGCTTAACACGACACAGAGAAGATCTATGATGACCAACCGCAAGTTCACCT 5616
5572 GCACCTTTGATCCGCTTAACACGACACAGAGAAGATCTATGATGACCAACCGCAAGTTCACCT 5631
5617 TCGGATTTCTGACACAGGCGGGCGCGCCAGCTCTGCTCAACCCAGCAGAGCTGAA 5676
5632 TCGGATTTCTGACACAGGCGGGCGCGCCAGCTCTGCTCAACCCAGCAGAGCTGAA 5691
5677 TGGTGTCAACGTGACATCTCCCTTGGGGGTACATTTGCTGGCATCCAGAGGCGCATCAT 5736
5692 TGGTGTCAACGTGACATCTCCCTTGGGGGTACATTTGCTGGCATCCAGAGGCGCATCAT 5751
5737 GTCTGAAAGAAATGGAATACGACAGCGGGCGCGCATCATCTCCAGGATCTTCTGCTGATGG 5796
5752 GTCTGAAAGAAATGGAATACGACAGCGGGCGCGCATCATCTCCAGGATCTTCTGCTGATGG 5811
5797 GAAGACATGGAGCTACATCTAGAG-----AGTCCATGTGTGCTACT 5844
5812 GAAGACATGGAGCTACATCTAGAGAGCGAGGTGTCCAGTCCATGGTGTGCTACT 5871
5845 ACACAGCCAGAGGAGTATCTTTGAGTTCGACAAGAAATCACCGCTCTCTTCTGTGAC 5904
5872 ACACAGCCAGAGGAGTATCTTTGAGTTCGACAAGAAATCACCGCTCTCTTCTGTGAC 5931
5905 GATGCCCAAGTGGCGCGGACACATGAGACCATCCGCTCAGTGGGCTACTACAGAA 5964
5932 GATGCCCAAGTGGCGCGGACACATGAGACCATCCGCTCAGTGGGCTACTACAGAA 5991
5965 CATCTATCAGCCCCCTGAGGCAATGCTCAGTATACAGGACTTCACTAGGATGGCA 6024
5992 CATCTATCAGCCCCCTGAGGCAATGCTCAGTATACAGGACTTCACTAGGATGGCA 6051

QY 6025 CCTCCTTACACCTTCTACCTGGGCACTGGCGGAGGGTGATATACAAAGTATGCGCAACT 6084
Db 6052 CCTCCTTACACCTTCTACCTGGGCACTGGCGGAGGGTGATATACAAAGTATGCGCAACT 6111
QY 6085 GTCAAAGCTGCGAGAGAGCGCTCTATGACACCAACAAAGTCAAGTTTCACTATGACGAGAC 6144
Db 6112 GTCAAAGCTGCGAGAGAGCGCTCTATGACACCAACAAAGTCAAGTTTCACTATGACGAGAC 6171
QY 6145 GCGAGGCACTGTAAGAACCTCAACCTACAGAAATGAGGGCTTCACTGACCATCCGCTA 6204
Db 6172 GCGAGGCACTGTAAGAACCTCAACCTACAGAAATGAGGGCTTCACTGACCATCCGCTA 6231
QY 6205 CCGTCAGATTTGGGCGCCCTGATTGACCGACAGATCTTCGGCTTCACTGAGGAAGGCAATGGT 6264
Db 6232 CCGTCAGATTTGGGCGCCCTGATTGACCGACAGATCTTCGGCTTCACTGAGGAAGGCAATGGT 6291
QY 6265 CAACGCCGTTTGAATAAATACTATGACAAACAGCTTCGGGTGACCAAGCATGCAAGGCTGT 6324
Db 6292 CAACGCCGTTTGAATAAATACTATGACAAACAGCTTCGGGTGACCAAGCATGCAAGGCTGT 6351
QY 6325 GATCAACGAGACCCACATGCCATTTGATCTCTATCGCTATGATGATGTCAGGCAAGAC 6384
Db 6352 GATCAACGAGACCCACATGCCATTTGATCTCTATCGCTATGATGATGTCAGGCAAGAC 6411
QY 6385 AGAGCAGTTTGGGAAGTTTGGTGTCAATTTACTATGACATTAACAGATCATFACCAACAGC 6444
Db 6412 AGAGCAGTTTGGGAAGTTTGGTGTCAATTTACTATGACATTAACAGATCATFACCAACAGC 6471
QY 6445 TGTATGACCCACACCAAGCAATTTGATGCAATATGGGAGGATGAGGAGTGCAGTATGA 6504
Db 6472 TGTATGACCCACACCAAGCAATTTGATGCAATATGGGAGGATGAGGAGTGCAGTATGA 6531
QY 6505 GATCTTCGGCTCGCTCATGTACTGATGACCGCTCAGTATGATAACATGAGGGCGAGTAGT 6564
Db 6532 GATCTTCGGCTCGCTCATGTACTGATGACCGCTCAGTATGATAACATGAGGGCGAGTAGT 6591
QY 6565 GAAGAAGGAGTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCTCTATGATGATGA 6624
Db 6592 GAAGAAGGAGTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCTCTATGATGATGA 6651
QY 6625 TGCTGACGCGCAGCTGACAGAGCTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTA 6684
Db 6652 TGCTGACGCGCAGCTGACAGAGCTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTA 6711
QY 6685 CGACCTCAATGGGAACCTTGCACTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACT 6744
Db 6712 CGACCTCAATGGGAACCTTGCACTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACT 6771
QY 6745 ACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAGATGATGA 6804
Db 6772 ACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAGATGATGA 6831
QY 6805 GGATGGCTTCTGAGCGACGCGGCGGTGATATCTTTGAGTACAACCTCAGCTGGCGCTGCT 6864
Db 6832 GGATGGCTTCTGAGCGACGCGGCGGTGATATCTTTGAGTACAACCTCAGCTGGCGCTGCT 6891
QY 6865 CATCAAGGCTTACAACCGGGCTGGCAGCTGGAGTGTACGGTACCGCTACGATGGCGCTGGG 6924
Db 6892 CATCAAGGCTTACAACCGGGCTGGCAGCTGGAGTGTACGGTACCGCTACGATGGCGCTGGG 6951
QY 6925 GCGGCGCTGTCACGACAGAGCAGCCACAGCCACCACTGCGAGTTCCTTATGACAGACCT 6984
Db 6952 GCGGCGCTGTCACGACAGAGCAGCCACAGCCACCACTGCGAGTTCCTTATGACAGACCT 7011
QY 6985 GACCAACCCCAACCAAGTCAACCCCTGTGTACAACCACTCCAGCTCTGAGATCACTCCCT 7044
Db 7012 GACCAACCCCAACCAAGTCAACCCCTGTGTACAACCACTCCAGCTCTGAGATCACTCCCT 7071
QY 7045 CTACTACGACTTTCAGAGACACTCTTTGCCATGAGAGCTGAGCAGTGTGATGATGATTTA 7104
Db 7072 CTACTACGACTTTCAGAGACACTCTTTGCCATGAGAGCTGAGCAGTGTGATGATGATTTA 7131
QY 7105 CATAGCTTGTGACAACTCGGACCCCTCTTCTGCTCTTTAGTGGAAACAGGTTTATGAT 7164

Db 7132 CATAGCTTGTGACAACTCGGACCCCTCTTCTGCTCTTTAGTGGAAACAGGTTTATGAT 7191
QY 7165 CAAAGCAAAATCTGTATACAGCCTATGGGGAGATCTATATGATATCAACACCCAACTTTCA 7224
Db 7192 CAAAGCAAAATCTGTATACAGCCTATGGGGAGATCTATATGATATCAACACCCAACTTTCA 7251
QY 7225 GATCATCATAGCTACCATGCTGGCTCTATCATCCACTCAACAGCTTGTCCACATGGG 7284
Db 7252 GATCATCATAGCTACCATGCTGGCTCTATGATATCACTCAACAGCTTGTCCACATGGG 7311
QY 7285 CCGGCGAGATTTATGATGTCTGGCGGACGCTGGACTAGCCAGACCAAGCTGTGGAA 7344
Db 7312 CCGGCGAGATTTATGATGTCTGGCGGACGCTGGACTAGCCAGACCAAGCTGTGGAA 7371
QY 7345 GCACCTTAGTAGCAGCAACGCTCATGCTTTTAACTCTATATGTTCAAAAACAAACCC 7404
Db 7372 GCACCTTAGTAGCAGCAACGCTCATGCTTTTAACTCTATATGTTCAAAAACAAACCC 7431
QY 7405 CATCAGCAAACTCCAGCAACATCAAGTGTCTTATGACAGATGTTAAACAGCTGGCTGTCA 7464
Db 7432 CATCAGCAAACTCCAGCAACATCAAGTGTCTTATGACAGATGTTAAACAGCTGGCTGTCA 7491
QY 7465 CTTTGGATTCAGCTACAAACGCTGATCCCTGGTTATCCAAACAGACATGATGCCAT 7524
Db 7492 CTTTGGATTCAGCTACAAACGCTGATCCCTGGTTATCCAAACAGACATGATGCCAT 7551
QY 7525 GGAACCCCTCTACGAGCTCATCCACACAGATGAAAACGAGGAGTGGACAAACAGCA - 7583
Db 7552 GGAACCCCTCTACGAGCTCATCCACACAGATGAAAACGAGGAGTGGACAAACAGCA 7611
QY 7584 -----AGTCTATCTCGGGGTACAGTGTGAAGTACAGAAAGCA 7620
Db 7612 GGTAAATCTGTGACAAAGGCTGCCAGTCTATCTCGGGGTACAGTGTGAAGTACAGAAAGCA 7671
QY 7621 GCTCAAGGCTTGTCACTTGAACGTTTGAACGCTCTATGGCTCCACATCAACAG 7680
Db 7672 GCTCAAGGCTTGTCACTTGAACGTTTGAACGCTCTATGGCTCCACATCAACAG 7731
QY 7681 CTGCCAGAGGCTCCAAAGCAAGAAAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGG 7740
Db 7732 CTGCCAGAGGCTCCAAAGCAAGAAAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGG 7791
QY 7741 GGTCAAGTTTGCCTTGAAGATGGCGAGTGCACACATCATCAGTGTGGCCATGA 7800
Db 7792 GGTCAAGTTTGCCTTGAAGATGGCGAGTGCACACATCATCAGTGTGGCCATGA 7851
QY 7801 GGATGGCGAAGGTTTGTGCCATCTTGAACCAATGCCACTACCTAGAGAACCTGCACCT 7860
Db 7852 GGATGGCGAAGGTTTGTGCCATCTTGAACCAATGCCACTACCTAGAGAACCTGCACCT 7911
QY 7861 CACCATTTGATGGGTGATACCCATTTACTTTGTGAAAACAGGACCTTCAGAGGTGACCT 7920
Db 7912 CACCATTTGATGGGTGATACCCATTTACTTTGTGAAAACAGGACCTTCAGAGGTGACCT 7971
QY 7921 GGCATCTCTGGGCTCAGTGGGGGGCGGAAACCTTGGAGATGGGCTCAACCTCACTGT 7980
Db 7972 GGCATCTCTGGGCTCAGTGGGGGGCGGAAACCTTGGAGATGGGCTCAACCTCACTGT 8031
QY 7981 GTCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGATCCAGTCCA 8040
Db 8032 GTCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGATCCAGTCCA 8091
QY 8041 GTACGGGGCACTGTGTTGAACACAGCTACGGGACAAAGTGGATGAGAGAGGACG 8100
Db 8092 GTACGGGGCACTGTGTTGAACACAGCTACGGGACAAAGTGGATGAGAGAGGACG 8151
QY 8101 GGTCTTGAGCTGGCGCGGACAGAGCCGTGCGCCCAAGCTGGCGCCCGCGAGCAGAG 8160
Db 8152 GGTCTTGAGCTGGCGCGGACAGAGCCGTGCGCCCAAGCTGGCGCCCGCGAGCAGAG 8211
QY 8161 ACTGCGGGAAGGGAAGGAGGCTTGGGGCTTGGACAGAGGGGGAAGCAGCAGGTGCT 8220

QY 542 CTGCGAAGACGAGCGCGCTCGGAGCGCGCGCGCGCTCTCGAGCGCCACACCC 601
Db 695 CTGCGAAGACCCCTCGGCTCGGAGCGCTCTCCACCACTGCCCGCCAGCCCATACCC 754
QY 602 AACAGACGACGCGCGCTCAATTAATCCCTGAACCGGGCAACTTCACGCCGAGGAGC 661
Db 755 AACAGACGACGCGCGCTCCATCAATCCCTTGAACAGGGGCACTTCACCCCAAGGAGC 814
QY 662 AACCCAGCGCGCGCCCGACGGAACCATCTGCTCTCCGGAGAGCGCCCTGCGGGCGGCC 721
Db 815 AACCCAGCGCGCGCCCGACGGAACCATCTCACTCTCTGCGGAGCGCCCGAGGAGGCC 874
QY 722 CAGGAGCCTGCGCGCGCCAGGAGAACTGGCTGTCTCAACGACCAATCCCTCGAGACC 781
Db 875 CAGGAGCAACCCATGCGCGCGAGCAAACTGGCTACTCAACAGTAATCCCACTGGAACC 934
QY 782 AGGAACCTAGGCAAGCAGCCATTCTAGGGAATTTGCAAGCAACCTCATTTAGATGGA 841
Db 935 AGAAACCTAGGCAAGCAGCCATTCTAGGGAATTTGCAAGCAACCTCATTTAGATGGA 994
QY 842 ATTCTCGGGGCTCCCGCATGATGGGCTTACAGTGACGGGCACTTCTCTCAAGCCT 901
Db 995 ATTCTCAGCGGCTCCCGCATGATGGGCTTACAGTGACGGGCACTTCTCTCAAGCCT 1054
QY 902 GGAGGCACTCCCGCTCTTCTGCAACACATACCAGGCTACCCACTGACCTCCAGCACA 961
Db 1055 GGAGGCACTCCCGCTCTTCTGCAACACATACCAGGCTACCCACTGACCTCCAGCACA 1114
QY 962 GTGTACTCTCTCCCGCGAGCCCTGCGCGAGCACTTCGCGCGCGCGCTTTAAAC 1021
Db 1115 GTGTACTCGCGCGCAACCGCGCGCTGCGCGAGCACTTCTCCCGAGCAGCCTTTAAAC 1174
QY 1022 CTCAGAAGACCTCAAGTACTGTAACTGGAAGTCGCGAGCCCTGAGCGCCATCGTCATC 1081
Db 1175 CTCAGAAGACCTTCCAGTACTGTAACTGGAAGTCGCGAGCGCTGAGCGCCATCTCATC 1234
QY 1082 TCAGGCACTCTGCTGTCATCTCTGTCATCTTGTGGCACTGCACTTGTGGCTTAAAC 1141
Db 1235 TCAGTACGCTGCTCATCTCTGTCATCTTGTGGCACTGCACTTGTGGCTTAAAC 1294
QY 1142 TGGCACTTCAGCCATGGAGG-----GCAGATGTATGATACAGGAGACACAGCC 1195
Db 1295 TGGCACTTCAGCCATGGAGGCGAGATGCAGATGTATGATACAGGAGACACAGCC 1354
QY 1196 AGCAGTTGGCTGTGCCAACCGACGCTCTCCCTATACCCCTCAGGGGCACTGGCTTAG 1255
Db 1355 AGCAGTTGGCTGTGCCAACCGACGCTCTCCCTGTATCCCTCAGGGGCACTGGCTTAG 1414
QY 1256 ACCCTGACAGGAAGGCAAGGAACCAAGGAAGGCGGAGGAGTCTTTCCAGAG 1315
Db 1415 ACCCTGACAGGAAGGCAAGGAGCGCGAGGAAGGCAAGGAGTCTTTCCAGAG 1474
QY 1316 GACAGTTTCATAGATTCGAGAAATGTATGGGAAGCGGAGCCTCCAGAAATTCCT 1375
Db 1475 GACAGTTTCATAGATTCGAGAAATGTATGGGAAGCGGAGTCTCCAGAAATTCG 1534
QY 1376 CTGCGACCTTCGAGATCTCAAGTGTTCATAGACCAATCTGTGCACTGAAATTCAT 1435
Db 1535 CTGGAACCTTCGAGATCTCAGGTGTTCATAGACCACTGACCCCTGTACCTGAATTCAT 1594
QY 1436 GTGTCTCTGGAAAGGCGCCCTGGTGGCAATTTATGGCAAGAAAGGCTCCCTCTTCA 1495
Db 1595 GTGTCTCTGGAAAGGCGCTCTGGTGGCAATTTATGGCAAGAAAGGCTCCCTCTTCC 1654
QY 1496 CATACACAGTTTGACTTTGTGAGTGTCTGGATGGCAGGAGCTCTTAACCCAGAGGCG 1555
Db 1655 CATACTCAGTTTGACTTTGTGAGTGTCTGGATGGGAAGAGGCTCTTAACCCAGAGGCA 1714
QY 1556 CGGAGCCTTAGAGGGAACCCCGCGCAGTCTCGGGAACTGTGCGCCCTTCAGCCCATGAG 1615
Db 1715 AGGAGCCTTGAGGGTCTCAGGCGCAATCAGCGGCGCTGTTCCTCCCGCATCCAGCCATGAG 1774

QY 1616 ACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGCACTTGTGCTTTTACAAATGACGA 1675
Db 1775 ACCGGTTTCATCCAGTATCTGGATTTCAGGAATCTGCACTTGTGCTTTTACAAATGATGA 1834
QY 1676 AAGGAGTCAGAAGTGTCTTCTTCTACCACTGCAATGAGTCGCTGGAATACCTGCCCC 1735
Db 1835 AAAGAATCTGAAGTGTCTTCTTCTACCACTGCAATGAGTCGCTGGAATACCTGCCCC 1894
QY 1736 AGCAATCTGATGGCAATGGTGAATCTGCACTTGTGCACTTGTGCTTCTGCGGTTTC 1795
Db 1895 AGCAATCTGATGGTGAATGGTGAATCTGCACTTGTGCACTTGTGCTTCTGCGGTTTC 1954
QY 1796 CTGGGCCCCGATGTGGCAGAGCCTCTGCCCCGCTCTGTAGCGGAAATGCGCCAAATAC 1855
Db 1955 CTGGGCCCCGATGTGGCAGAGCCTCTGCCCCGCTCTGTAGCGGAAATGCGCCAAATAC 2014
QY 1856 ATGAAGAGGAGATGCTTGTGCCACAGTGGCTGGAAGGCGCTGAGTCGATGTGCCAC 1915
Db 2015 ATGAAGAGGAGTGGCTGTGCCACAGTGGCTGGAAGGCGCTGAGTCGATGTGCCAC 2074
QY 1916 AACCACTGTATCGATGTGGCTGCAACCACTGCACTGCACTCAACGCGCACTTCATC 1975
Db 2075 AACCACTGTATCGATGTGGCTGCAACCACTGCACTGCACTCAACGCGCACTTCATC 2134
QY 1976 TGCAACCTCTGGTATCAAGGCGAGAGCTGTGAGGAAGTGGATGCAATGAGCCCAATGT 2035
Db 2135 TGCAACCTCTGGTATCAAGGCGAGAGCTGTGAGGAAGTGGATGCAATGAGCCCAATGT 2194
QY 2036 TCAGGCGGGGCTGTGCGTGAAGGAGATGCAATGCTTGTGGGATGGGAGGAGCACC 2095
Db 2195 TCAGGAGGGGCTGTGCTGAGAGGAGTGGCTGTGCTGTGGGATGGGAGGAGCACC 2254
QY 2096 AACTGCGAGACCCCGAGGCGCACATCTTAGACCACTGTTCAGGCCACCGAACTTCCTC 2155
Db 2255 AACTGCGAGACCCCGAGGCGCACATCTTAGACCACTGTTCAGGCCACCGAACTTCCTC 2314
QY 2156 CCGGACACCGGGCTTTGAGCTGTGACCCAGCTGAGTGGACACCACTGCTTATTCGAG 2215
Db 2315 CCGGACACAGGACTTTGCACTGTGACCCAGCTGAGTGGACACCACTGCTTATTCGAG 2374
QY 2216 ATCTGTCTCCGACTGTGTGGCCATGCGTGTGCTAGGGGCGCACTGCGGCTGCGAG 2275
Db 2375 ATCTGTCTCCGACTGTGTGGCCATGCGTGTGCTAGGGGCGCACTGCGGCTGCGAG 2434
QY 2276 GATGGCTGATGGGCGAGCTCGCAACGCGGCGCTGCCACCCGCGCTGTGCGGAGCAT 2335
Db 2435 GATGGTTGATGGGCGCGCATCGCAACGCGGCGCTGCCACCCGCGCTGTGCGAGAAC 2494
QY 2336 GGGACCTGCCCGACCGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACTTCGACC 2395
Db 2495 GGGACCTGCCCGGACCGCAAGTGCAGTGCAGCCCGGCTGGATGGAGAGCATGCAACC 2554
QY 2396 ATCGCTCATTCTGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGAAC 2455
Db 2555 ATCGCTCATTCTGATAGGGTAGTTAAAGAGGGCTGTCTGGCTGTGCAATGGAAT 2614
QY 2456 GCGAGATGATCTTAGACCTGAATGTTGGCACTGCTGCGAGCTGGGCTGGGCTGAGAG 2515
Db 2615 GCGAGATGATCTTAGACCTGAATGTTGGGCTGCTGCGAGCTGGGCTGGGCTGAGAG 2674
QY 2516 GCTGGCTGTGCACTTCCATGAGACTGCTGCGGCTGACAGCAAGAACTGATGAGAT 2575
Db 2675 ACTGGCTGCGACACATCCATGGAAACGGGCTGTGGAGCGGCAAGGACACGCGAGAT 2734
QY 2576 GGGCTGTGAGTGTGACCTGTGACCCCTGTGCTGCTGCGAGCCCTGTGCGCATATCAACCCG 2635
Db 2735 GGGCTGTGAGTGTGACCTGTGACCCCTGTGCTGCTGCGAGCCCTGTGCGCATATCAACCCG 2794
QY 2636 CTGTGCTTGGCTCCCTTACCTCTGAGCATCATCCAGGAGACACAGGTCCTGTGTCA 2695
Db 2795 CTGTGCTTGGCTCCCTTACCTCTGAGCATCATCCAGGAGACACAGGTCCTGTGTATCC 2854
QY 2696 CAGCAGAACCTTACATCTCTTATGACCCGATCAAGTTCTCTGTGGGCGAGGACAGCAGC 2755

[illegible]

QY 4916 ATGGTAAATGTCGCGCAGAGCTCTACTGGATGCGCTCTGGCTGGTGGTCCAGATGGC 4975
 Db 5075 ATGGTGAACGTCGCGCAGAGCTCTACTGGATGCGCTCTGGCTGGTGGTCCAGATGGC 5134
 QY 4976 CAGGTCTACTGGGTGACCATGGGACACCAAGTGCATCAAGAGTGTGACACACACAGGA 5035
 Db 5135 CAGGTATATCTGGGTAAACCATGGGACACCAAGTGCATCAAGAGTGTGACACACACAGGA 5194
 QY 5036 CACGAGTTGGCCATGATGACATACCATGGCAATTCGCGCTTCCTGGCAACCAAAAGCAAT 5095
 Db 5195 CACGAGCTAGCCATGATGACCTACCATGGCAATTCGCGCTTCCTGGCAACCAAAAGCAAT 5254
 QY 5096 GAAACCGATGACCAACATTTATGAGTACGACAGCTTTGGCCCTCGACCAATGTGACC 5155
 Db 5255 GAAACCGGTGACCAACATTTATGAGTACGACAGCTTTGGCCCTCGACCAACGTGACC 5314
 QY 5156 TTCCCTACTGGCCAGGTGAGCAGTTTCGGAAGTATACAGACAGTTCAAGTGCATGTCCAG 5215
 Db 5315 TTTCCTACTGGCCAGGTGAGCAGTTTCGGAAGTATACAGACAGTTCAAGTGCATGTCCAG 5374
 QY 5216 GTAGAGACTCGAAGAGATGTACACATACCAACCAACCTGTCTGCGCTCAGGCGCC 5275
 Db 5375 GTAGAGACTCGAAGAGATGTACACATACCAACCAACCTGTCTGCTTCGGGTGCC 5434
 QY 5276 TTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGCCGAGTCC 5335
 Db 5435 TTCTACACCTGTTTCAAGACCAAGTCCGGAACAGCTACTACATCGGGCCGAGTCC 5494
 QY 5336 TTGCGGCTGCTGCTGGCCAAACGGCATGGAGTGGCGCTGCAGACTGAGCCCACTTGGT 5395
 Db 5495 CTGAGGCTGCTACTAGCCAAATGGCATGGAAGTGGCTCTGCGAGCTGAGCCACACTGGT 5554
 QY 5396 GCTGCGACCTCAACCCACCGTGGCGAAGAGGATGTCAAGTCCGCTCCCATCGAACCGC 5455
 Db 5555 GCTGCGACTGTCAACCCCACTGTAGGCAAGAGGATGTCAAGTCCGCTCCCATCGAACCGC 5614
 QY 5456 CTCACCTGCTGGAGTGGCGCCAGCGCAAGAGAGGCTCGGGCCAGGTCACCTGTCTTT 5515
 Db 5615 CTCACCTGCTGGAGTGGCGCCAGCGCAAGAGAGGCTCGGGCCAGGTCACCTGTCTTT 5674
 QY 5516 GGGCGCGGCTCGGGTGACAAACCGAAATCTCTATCTCTGGAATTTGATCGGGTAAACA 5575
 Db 5675 GGACGCGTCTCGGGTTTCAACCGAAACCTCTTGTCTTTGGAATTTGACCGGTGTACA 5734
 QY 5576 CGCAGAGAGAGATCTATGATGACCAACCGAGTTTCAACCTTCGGATCTGTAGACACAG 5635
 Db 5735 CGCAGAGAGAGATCTATGATGACCAACCGAGTTTCAACCTTCGGATCTGTATGACACAG 5794
 QY 5636 GCGGGCGCGGCTCTGGTCAACCGACAGAGGCTGAATGGTGTCAACGTGACATAC 5695
 Db 5795 GCAGGGAGGCCAGGCTCTGGTCACTAGCAGAGGCTGAATGGTGTGAATGTGACCTAC 5854
 QY 5696 TCCCTGGGGTTTACATTTCTGGCATCCAGAGGGGATCATGTCTGAAGATGAATAC 5755
 Db 5855 TCCCTGGAGGTACATTTCTGGATTCGAAGGGGATCATGTCTGAGAGATGAATAT 5914
 QY 5756 GACGAGCGGGCGGATCAATCCAGGATCTTCTGGTGGGAAGACATGGAGTGTACACA 5815
 Db 5915 GATCAGCAGGGGATCAATCCCGGATCTTTCAGAGCGGAAATGTGGAGCTACAG 5974
 QY 5816 TACTTAGAAGTCCATGGTGTGTATACACAGCAGAGGACGTATATCTTTGAGTTC 5875
 Db 5975 TACTTAGAAGTCCATGGTGTGTATCTTCCACAGCAGAGGACGTATATCTTCCAGTTC 6034
 QY 5876 GACAAATGACCGCTCTCTCTGTGAGGATGCCAACGTGGCGGCGAGACACTAGAG 5935
 Db 6035 GACAAATGACCGCTCTCTCTGTGAGGATGCCAACGTGGCGGCGAGACCGTGGAG 6094
 QY 5936 ACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCTCAGGGCAATCCCTCA 5995
 Db 6095 ACCATCCGCTCAGTGGGCTACTACAGAAACATCTACAGCCCTCAGGGCAATCCCTCA 6154

QY 5996 GTCATACAGGACTTCACTGAGGATGGGACCTCTTTCACACCTTCTACTGGGACTGGC 6055
 Db 6155 GTCAITCAGGACTTCACTGAGGATGGACACCTGTTCATACCTTCTACTTGGGACCGGC 6214
 QY 6056 CGCAGGCTGATATACAAAGTATGGCAAACTGTCAAAGCTGGCAGAGCGCTCTATGACACC 6115
 Db 6215 CGCGGGTGTATTTACAAAGTATGGCAAGTGTCAAAGCTGGCAGAGCTCTGTATGACACC 6274
 QY 6116 ACCAAGGTCACTTTTCACTTATGACGAGAGCGGAGGATGCTGAAGCAATCAACCTACAG 6175
 Db 6275 ACTAAGGTAACTTCACTTACGAGAGCGGAGGATGCTGAAGCAATCAACCTACAG 6334
 QY 6176 AATGAGGGCTTCACTTACCTGACCAATTCGCTACCGTCAAGTTGGGCCCTGATTTGACCGACAG 6235
 Db 6335 AATGAGGGCTTCACTTACCTGACCAATTCGCTACCGTCAAGTTGGGCCCTGATTTGATCGGACAG 6394
 QY 6236 ATCTTCCGCTTCACTTACGAGGAGGATGTTCAACCGCCGCTTTTGACTACAACTATGACAAC 6295
 Db 6395 ATCTTCCGCTTCACTTACGAGGAGGATGTTCAACCGCCGCTTTTGAATTTACAACTATGACAAC 6454
 QY 6296 AGCTTCCGGGTGACGAGCATGAGGCTGTGATCAACGAGACCCACTGCGCCATTTGATCTC 6355
 Db 6455 AGTTTCCGCTGACTAGCATGAGGCTGTGATCAATGAGACCCCACTGCGCCATTTGACCTC 6514
 QY 6356 TATCGCTATGATGTGTGAGGAAAGACAGAGCAGTTTGGGAAGTTTGGTGTCAATTTAC 6415
 Db 6515 TACCGCTATGATGTGTGAGGAAAGACAGAGCAGTTTGGGAAGTTTGGTGTCAATTTAC 6574
 QY 6416 TATGACATTTAAACAGATCATCACAGCTGTGTATGACCCACACCAAGCATTTGATGCA 6475
 Db 6575 TACGACATCAACAGATCATTTACCAAGCGCTCATGACCCACACCAAGCATTTGATGCT 6534
 QY 6476 TATGGCAGGATGAAGAAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATGACC 6535
 Db 6535 TATGGCAGGATGAAGAAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATGACT 6594
 QY 6536 GTCCAGTATGATTAACATGGGGGAGTGTGAAGAGAGAGCTGAAGTAGGACCTTAGCC 6595
 Db 6695 GTTCAGTATGATTAACATGGGAGGCTGTGAAGAGAGAGCTGAAGTAGGACCTTAGCC 6754
 QY 6596 AATACCACTCGCTACTCTCTATGATGTGATGTGATGAGTCTTCCGCTCGCTCATGTACTCCATC 6655
 Db 6755 AACACTACCGCTACTCTCTATGATGTGATGTGATGAGTCTTCCGCTCGCTCATGTACTCCATC 6814
 QY 6656 AATGACAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCACTTACTGAGC 6715
 Db 6815 AATGACAAGCCACTCTGGCGCTACAGCTATGACCTCAATGGGAACCTTACACTTGTGAGC 6874
 QY 6716 CCTGGGAACAGTGACCGGCTCACACCACTAGGATGATGATCCGCGACCGCATCACTCGG 6775
 Db 6875 CCTGGGAACAGTGACCGGCTCACACCACTAGGATGATGATCCGCGACCGCATCACTAGG 6934
 QY 6776 CTGGGTGACGTGCAATAAAGATGGATGGATGGCTTCTGAGGAGCGGGGGGGTGTAT 6835
 Db 6935 CTGGGTGATGTACAGTACAGATGGATGGATGGCTTCTGAGGAGCGGGGGTGGGAT 6994
 QY 6836 ATCTTTGAGTACAACTCAGCTGGCTGTCTCATCAAGGCTTACAAACCGGGCTGGCAGCTGG 6895
 Db 6995 GTCTTCAGTACAACTCAGCGGGCTGTCTCATCAAGGCTTACAAACCGGGCTGGGCTGG 7054
 QY 6896 AGTGTCAAGTACCGCTACGATGGCTGGGCGCGCTGTCCAGCAAGAGAGCAGCAAGC 6955
 Db 7055 AGTGTCAAGTACCGCTACGATGGCTGGGAGCGCGGATGTCCAGCAGAGAGCAGCAAGC 7114
 QY 6956 CACCACCTGAGTCTTCTATGAGAGCTTGAACCAACCCCAAGGTCAACCACTGTATC 7015
 Db 7115 CACCACCTGAGTCTTCTATGAGAGCTTGAACCAACCCCAAGGTCAACCACTGTATC 7174
 QY 7016 AACCACTCAGCTTGTAGATCACTCCCTCTACTACGACTTGAAGGACACCTCTTTGCC 7075
 Db 7175 AACCACTCAGCTTGTAGATCACTCCCTCTACTATGATCTGCAAGGACACCTCTTTGCC 7234
 QY 7076 ATGAGCTGAGCAGTGGTGTGATAGTTTTATACATAGCTTTGTGACAACTCGGGACCCCTCTT 7135

		misc_feature		1813 .. 1881		/gene="Duc4"		/note="putative; transmembrane-region site"			
		ORIGIN									
		Query Match		79.0%; Score 6601.2; DB 10; Length 9722;							
		Best Local Similarity		86.7%; Pred. No. 0;							
		Matches 7426; Conservative		5; Mismatches 874; Indels 264; Gaps 5;							
Qy	2	TTTGTGATGTGAGAGCGCGCGCGAGGCCATGGACGTGAGGAGGAGAGGAGCCCTTAC	61								
Db	550	TTCTGTGATGTGGAAGACGAGAGCCAGAGGCCATGGACGTGAGGAGGAGAGCCCTAT	609								
Qy	62	CGCTCGCTGACCCGCGCGCGAGCGCGCGCGCTACACAGCTCGTCGCGCGGACAGC	121								
Db	610	CGCTCGCTGACCCGCGCTCGCGATGCTGAGCGCGCGCTACACAGCTCATCGGACAGC	669								
Qy	122	GAGGAGGCAAGCCCGCAGAAATCGTACAGCTCCAGCGAGACCTTGAGGCGCTACGAC	181								
Db	670	GAGGAGGCAAGGGCCACAGAGTCTCTACAGCTCCAGTGTAGACCTCTGAAGGCCATATGAC	729								
Qy	182	CAGGACCCCGCTAGCCTATGGCAGCGCGCTCAGGACATTTGTCGCGCAGAGGCGCGAG	241								
Db	730	CAAGATGCTCGCCTAGCCTACGGCAGCGCGCTCAGGACATGTTACACAGGAGCGCGAG	789								
Qy	242	CAATTCTCGCGCACAGGTGCCAACTTCACCCCTGCGGGAGCTGGGGCTGGAAAGATGACG	301								
Db	790	GAGTTCGCGCACAGGCACTAAATTTCACTCTCTGTAAGTGGGACTGGGAGAGATGACG	849								
Qy	302	CCCCCTACGGGACCTGTATCGGACAGACATTTGGCCTTGCCCCAATGCGGCTATCTCCATG	361								
Db	850	CCCCCTCATGGGACTCTCTACAGGACAGACATCGGCCTCCCACTGTGGGTATTTCCATG	909								
Qy	362	GGGGTGGCTCTGATGCGGACATGGAGGCTGACACGGTGTCTGCCCTGAGCACCCCGGTG	421								
Db	910	GGGGCCAGCTCTGATGTCAGACTTGGAGACAGACACTGTGTCTATCCCTGAACACCCAGTG	969								
Qy	422	CGTCTGTGGGCGGAGCACACGTTACGGGCGCAGCTCCTGTGCTGTCAGCGCGGCCAAT	481								
Db	970	CGCCTGTGGGCGGAGCACACGTTACGGGCGCAGTTCTTCGCTGTCCAGCGCGGCCAAC	1029								

1346 GTGGGAAGCGAGCTCCAGAGATTCTCTCGGACCTTTCTGGAGATCTCAAGTGTC 1405
1406 ATAGACCACTCTGTGCATCTGAATTTCAATGTGTCTCTGGGAAAGCGAGCCCTGTGGC 1465
1466 ATTTATGGCAGAAAGGCTCCCTCTCTTACATACACAGTTTGACTTTGTGGAGTGTG 1525
2230 ATTTATGGCAGAAAGGCTCCCTCTCTTCCCATCTAGTTAGACTTTGTGGAGTCTGT 2289
1526 GATGGCAGAGAGCTCTTAACCCAGAGGCGGAGCCTTAGAGGGAGACCCCGCGCAGTCT 1585
2290 GATGGAAGAAGCTCTTAACCCAGAGGCGGAGCCTTAGAGGGTCTTACGGGCCATCA 2349
1586 CCGGGAACTGTGCCCTCTCCAGCCATGAGACAGGCTTCAATCCAGTATTTGGATTAGGA 1645
2350 CCGGGCCCTGTTCCTCCATCCAGCCATGAGACCGGTTTCAATCCAGTATCTGGATTAGGA 2409
1646 ATCTGGCACTTGGCTTTTACAAATGACGGAAGGAGTCAAGAGTGGTTTCTTCTCACC 1705
2410 ATCTGGCACTTGGCTTTTACAAATGAGTGAAGAAGAACTGAAGTGGTTTCAATTTCTCACC 2469
1706 ACTGCCATGAGTCGGTGATTAATGCCCCAGCAACTCTATGGCAATGGTACTGCATC 1765
2470 ACCGCCATGGAATCAGTGGATAAATGCCCCAGCACTGTAAGAGGCGAGTGGCCTGCAAT 2529
1766 TCTGGAGCTGCATGCTTCTGGGTTTCTGGGCTTCTGGGCTTCTGGGCTTCTGGGCTTCTGGG 1825
2530 TCGGGACCTGCTGCTTCTGGGTTTCTGGGCTTCTGGGCTTCTGGGCTTCTGGGCTTCTGGG 2589
1826 CCGCTGCTGTAGCGGAAATGGCCAAATACATGAAAGGAGAGTGTGTGCGCAAGTGGC 1885
2590 CCGCTGCTGTAGCGGAAATGGCCAAATACATGAAAGGAGAGTGTGTGCGCAAGTGGC 2649
1886 TGGAAAGGCTGAGTGGATGTCGCCAACCAACGATGTATCGATGTGGCTGCGGCTGAGCAAC 1945
2650 TGGAAAGGCTGAGTGGATGTCGCCAACCAACGATGTATCGATGTGGCTGCGGCTGAGCAAC 2709
1946 CATGCCACTGTCATCGGCACTCTGATCTGCAACCTGCTCAAGGGGCGAGCTGT 2005
2710 CATGCCACTGTCATCGGCACTCTGATCTGCAACCTGCTCAAGGGGCGAGCTGT 2769
2006 GAGGAAGTGGATGTCATGGAACCCACATGTTTCAAGCGGGGTGTCTGGTGGAGGGCAA 2065
2770 GAGGAAGTGGATGTCATGGAACCCACATGTTTCAAGCGGGGTGTCTGGTGGAGGGCAA 2829
2066 TGCCATCTGTTGTGGATGGGAGGACCACTGCGAGACCCCGAGGCGACATGCTTA 2125
2830 TGCCATCTGTTGTGGATGGGAGGACCACTGCGAGACCCCGAGGCGACATGCTTG 2889
2126 GACCACTGTTGAGCCACCGAACCTTCTCCCGGACACCGGGCTTTCAGCTGTGACCCA 2185
2890 GACCACTGTTGAGCCACCGAACCTTCTCCCGGACACCGGGCTTTCAGCTGTGACCCA 2949
2186 AGCTGGATGGACAGGCTTCTTATCGAGATCTGCTGCGGCTGCTGCTGCGGCTGCTG 2245
2950 AGCTGGATGGACAGGCTTCTTATCGAGATCTGCTGCGGCTGCTGCTGCGGCTGCTG 3009
2246 GTGTGCTGAGGGGCACTGCGGCTGCGAGATGGCTGGATGGGGGAGCGCTGCGACAG 2305
3010 GTGTGCTGAGGGGCACTGCGGCTGCGAGATGGCTGGATGGGGGAGCGCTGCGACAG 3069
2306 CCGGCTTCCACCGGCTGTCGAGCATGGGACCTGCGGCTGCGGCTGCGGCTGCGGCTG 2365
3070 CCGGCTTCCACCGGCTGTCGAGAAACACGGGACCTTCCCGGAGCGGCAAGTGGATGC 3129
2366 AGCCCTGCTGGAATGGGAAACACTGCAACCATCTGCTCACTATCTGGATAGGGTAA 2425
3130 ACCCCGGCTGGAATGGAGACACTGCAACCATCTGCTCACTATCTGGATAGGGTAA 3162
2426 GAGGTTGCCCTGGGTTGTGCAATGGGCAACGGCAGATGTACCTTAGACCTGAATGGTTGG 2485

3163 GAGGTTGCCCTGGGTTGTGCAATGGAAATGGCAGATGTACCTTGACCTGAATGGTTGG 3222
2486 CACTGGCTGTGCCAGCTGGGCTGGAGAGAGTGGCTGTGACACTTCCATGGAGACTGCC 2545
3223 CACTGTGTCTGCCAGCTGGGCTGGGAGGAGTGGCTGCGACACATCCATGGAACGGGC 3282
2546 TGGGTTGACAGCAAAACAAATGATGGAGATGGCTGGTGGAGTGCATGACCTGACTGC 2605
3283 TGTGGAGACGGCAAGCAACACGAGATGGCTGGTGGAGTGCATGACCTGACTGC 3342
2606 TGCTCCAGCCCTGTGCCATATCAACCCCTGTGCTTGGCTGCCCTAACCTCTGGAC 2665
3343 TGCTCCAGCCCTGTGTCTGATGTCAACCCCTGTGCTTGGCTGCCCTGACCCCTGAC 3402
2666 ATCATCCAGAGACACAGCTTCCCTGTGTCAAGCAGAACCTACACTCTTCTATGACCGC 2725
3403 ATCATCCAGAGACACAGCCCTCTGTATCCAGCAGAACCTGAACTCTTCTACGATCGA 3462
2726 ATCAAGTTCTCTGTGGGAGGACAGCAGCAGATAATCCCGGGAGAACCCCTTTGAT 2785
3463 ATCAAGTTCTCTGTGGGAGGACAGCAGCAGATAATCCCGGGAGAACCCCTTTGAC 3522
2786 GGAGGCACTGCTGTGTATTTCTGTGCCCAAGTGTGATGATCATGATGGAACCCCTTGGT 2845
3523 GGAGGCACTGCTGTGTATTTCTGTGCCCAAGTGTGATGATGATGGAACCCCTTGGT 3582
2846 GGTGTGAACATCAGTTTGTGCAATTAACCTCTCTTTGGATATACATCAGCAGCAAGAT 2905
3583 GCGTGAACATCAGTTTGTGCAATTAACCTCTCTTTGGATATACATCAGCAGCAAGAT 3642
2906 GGAGGCTTTGACTTGTGCAAAATGGCGGCACTCTCCATCATCTCGGCTTTCAGCGGCA 2965
3643 GGAGGCTTTGACTTGTGCAAAATGGCGGCACTCTCCATCATCTCGGCTTTCAGCGGCA 3702
2966 CTTTTCATCAGCAGGACACACCTGTGGCTGCCATGGGATCGCTTCTTTGTCTAGGAA 3025
3703 CCGTTTCATCAGCAGGACACACCTGTGGCTGCCATGGGATCGCTTCTTTGTCTAGGAA 3762
3026 ACCATCATCAGCAGGACACACCTGTGGCTGCCATGGGATCGCTTCTTTGTCTAGGAA 3085
3763 ACCATGTCTAGCAGCAGGAGATGAGATCCCGCAGCAGACCTGAGCACTTTGCC 3822
3086 CCGCCCAACCCAGTGTCTCTCCATCCCACTGAGCTCTTTCGCGAGCTCTCTGTCAGAG 3145
3823 CCGTCCCAACCCAGTGTCTCTCCATCCCACTGAGCTCTTTCGCGAGCTCTCTGTCAGAG 3882
3146 AAGGCCCCATTTGCCGMAATTCAGGCTTTCAGGAGGAAATCTCTATCTCTGGCTGC 3205
3883 AAGGCCCCATTTGCCGMAATTCAGGCTTTCAGGAGGAAATCTCTATCTCTGGCTGC 3942
3206 AAGATGAGCTGAGCTACCTGAGCAGCGGACCCCTGGCTACAAATCTGTCTGAGGATC 3265
3943 AAGATGAGCTTCTGAGCTACCTGAGCAGCGGACCCCTGGCTACAAATCTGTCTGAGGATC 4002
3266 AGCTTCCACCCAGCAGCTCTCAACAGAGGTTGGGCTTTCAGAGGCTTTCAGAGGCTTTC 3325
4003 AGTCTCACACACCCAGCAGCTCTCAACAGAGGTTGGGCTTTCAGAGGCTTTCAGAGGCTTTC 4062
3326 GAGGCCCCCTTTCAGGAAAGTGGTTCGCTGCGAGCCCCAGACCTGTCTCTATTTTCAAT 3385
4063 GAGGTTGGCTTCTCAGGAAAGTGGTTCGCGCAGCTCTGACCTGTCTTACTTCTATC 4122
3386 TGGGACAGACAGAGCTCTCAACAGAGGTTGGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGGCTTTC 3445
4123 TGGGACAGACAGAGTCTCAACAGAGGTTGGGCTTTCAGAGGCTTTCAGAGGCTTTCAGAGGCTTTC 4182
3446 GTGGGTTATGAATGAATCTGCGCCAGATCTAATCTCTGCGGAAAGAAAGCAACAGTGC 3505
4183 GTGGGTTATGATGAATCTGCGCCAGATCTGATCTCTGCGGAAAGAAAGCAACAGTGC 4242
3506 CTGCAAGGCTATGAATGAATGAGCGGCTTCAAGCTTGGAGATGAGGCTTAGACAAATCATCAT 3565

Db	4243	CTGCAGGGCTATGAAATGATGCTCCAAAGCTGGGGGGCTGGAGTCTGGATAGACACCAT	4302
Qy	3566	GCCCTCAACATTCAAGATGGTATCCTGCACAAAGGGAATGGGAGAACCAAGTTTGTGTCT	3625
Db	4303	GCCCTGAACATCCAGAGTGGCATCTGCACAAAGGGAATGGAGAGAACCAAGTTTGTGTCT	4362
Qy	3626	CAGCAGCCTCTGTCTATTTGGAGGATCATATGGGCAATGGGGCCCGGAGAAAGCATCTCTGC	3685
Db	4363	CAGCAGCCACCGGTCTATGGGAGCATCATGGGCAATGGGCAATGGGCAAGATATCTCTCTGC	4422
Qy	3686	CCAGCTCAACAGCGCTTGTGACGGCAACAAAGCTCTGGGCCCGGAGTGGCCCTCACCTGT	3745
Db	4423	CCCAGCTGCATGTCTTGTGTATGGCAACAGCTCTGGGCCCGGAGTGGCCCTCACCTGC	4482
Qy	3746	GGCTCTGACGGAGCCTCTATGTGGGTGATTTTAACTACATTAGAGATCTTCCCTCT	3805
Db	4483	GGCTCTGATGGAGTCTCTAGTGGAGACTTCAACTACATCCGAGAACTTCCCTCT	4542
Qy	3806	GGAAATGTCACCAATCCTCGAGATG-----AGTCACAGCCCA	3865
Db	4543	GGAAATGTCACCAATCCTCGAGATG-----AGTCACAGCCCA	4581
Qy	3866	GCACAAATATCTACCTGGCCACAGACCCCATGATGGGGCCGTCTTCTCTTTGACAGC	3925
Db	4582	GCACAAATATCTACCTGGCTACAGACCCCATGATGGGGCCGTCTTCTCTTTGACACC	4641
Qy	3926	AACAGCCCGCGGTCTTTAAATCAAGTCCACTGTGTGTGTAAGGACCTTGTCAAGAAC	3985
Db	4642	AACAGCCCGCGGTCTTTCAAGGTCAAAGTCCACCAAGTGTGAAGGACCTTGTGAAGAAC	4701
Qy	3986	TCTGAGGTGGTTGGGGGACAGGTGACCAAGTGCCTCCCTTTGATGACACTCGCTGGCGG	4045
Db	4702	TCCGAGGTGGTAGCAGGACTGGTGACAGTGCCTCCCTTTGATATACCCGCTGGGA	4761
Qy	4046	GATGTGGGAAGGCCACAGAGCCACACTCCAAATCCAGGGGTATTACAGTGGACAAG	4105
Db	4762	GATGTGGGAAGGCCACAGAGCCACACTCCAAATCCAGGGGTATTACAGTGGACAAG	4821
Qy	4106	TTTGGGCTGATCTACTCTGTGATGGCACCACTGATCAGACCATCGATCAGAAATGGGATC	4165
Db	4822	TTTGGGCTGATCTACTCTGTGATGGCACCACTGATCAGACCATCGATCAGAAATGGGATC	4881
Qy	4166	ATCTCCACCTGCTCGGCTTAATGATCTCATCATCAGCCCGGCCACTCAGTGTGATCT	4225
Db	4882	ATCTCCACTTTGCTGGGTCCAATGACCTCACTCGGCGGCGCCCTCAGCTGTGACTCC	4941
Qy	4226	GTCATGGATATTTCCAGGTAAAGTGGAGTGGGCCACAGACTTAGCCATCAACCCCATG	4285
Db	4942	GTCAATGGAGATTTCTCAGGTTGGCTGGAGTGGGCCACAGACTTAGCCATCAACCCCATG	5001
Qy	4286	GACAACTCACTTTATGTCTCGACAACTATGTGTCTCTGCAAAATCTCTGAAACACCCAG	4345
Db	5002	GATAATTTCTCTATGTCTCGACAACTATGTGTCTCTGCAAAATCTCTGAAACACCCAG	5061
Qy	4346	GTGCGCATTTGCGCGGAGGCCCATGTACATGCGCAGGTCCCTGGCATTGACACTTCTGT	4405
Db	5062	GTACGCATTTGCTGGGAGGCCCATGTACATGCGCAGGTCCCTGGCATTGACACTTCTGT	5121
Qy	4406	CTAAGCAAGTGGCCATCCACGCAACCTGGAGTCAGCCACCGCTTTTGGCTGTTTCAAC	4465
Db	5122	CTGAGCAAGTGAACATCCATCGCACCTTGAGTGAAGCACTGCTCTGCGCGTTTCAAC	5181
Qy	4466	AATGGGGTCTCTATTTGCTGAGACTGATGAGAAAAAGATCAACCCCATCAGGAGTCT	4525
Db	5182	AATGGAGTCTTGTACATCGCTGAGACCGACGAGAAAAAGATCAATCCCATCAGGAGTCT	5241
Qy	4526	ACCACTAGTGGAGAGATCTCACTCTGCTGGGCGCCCGAGTGGCTGTGACTGTATAAAT	4585
Db	5242	ACTAAGTGGTGAATCTCACTGTTGCTGGTGGCGCCCGAGTGGCTGTGACTGTATAAAT	5301
Qy	4586	GATGCCAAGTGTGATTTTCTGGAGACGATGTTATGCAAGGATGCAAGTAAAT	4645
Db	5302	GATGCCAAGTGTGATTTTCTGAGATGATGTTAGGCCAAGGATGCAAGTAAAT	5361
Qy	4646	ACCCATCTCTTGGCTGTGTGTGATGGGGAGCTCTACCTGGCCGACCTTGGGAC	4705
Db	5362	ACCCATCTCTTGGCTGTGTGTGATGGGGAGCTCTATGTGGCCGACCTTGGGAAAC	5421
Qy	4706	ATCCGAATTCGGTTTATCCGGAAGAACAGCTTCTTCCCTCAACACCCAGAACATGTATGAG	4765
Db	5422	ATCCGAATTCGAATTTATCCGGAAGAACAGCTTCTTCCCTGAAACACTCAGAACATGTACGAG	5481
Qy	4766	CTGTCTTCCACCAATTGACACAGAGCTTATCTTGTGTATACCCCGGCAAGCACTGTATAC	4825
Db	5482	CTATCTCTCCCATCGACAGGAGCTGTACCTCTTTGATACCACTGGCAAGCATCTGTAC	5541
Qy	4826	ACCCAAAGCTGCCACAGAGACTACCTGTACACTTCACTACACTGAGGAGCGGCGAC	4885
Db	5542	ACTCAGAGCTTACCCACAGGAGACTACCTGTACACTTCACTTACAGGGAGCGGCGAC	5601
Qy	4886	ATCACACTCATCACAGACAAATGGCAACATGGTAAATGTCCCGGAGACTCTACTGGG	4945
Db	5602	ATCACACATATCACGACAAATGGCAACATGGTAAATGTCCCGGAGACTCTACTGGG	5661
Qy	4946	ATGCCCTCTGGCTGGTGTCCAGATGGCCAGGTGTACTGGGTGACCATGGGCGCCACAC	5005
Db	5662	ATGCCCTCTGGCTGGTGTCCAGATGGCCAGGTGTACTGGGTAACTGGGCGCCACAC	5721
Qy	5006	AGTGCACTCAAGAGTGTGACACACAAAGGACACGAGTTGGCCATGATGACATACCATGGC	5065
Db	5722	AGCGCACTCAGAAGTGTGACACACAAAGGACACGAGTAGCCATGATGACCTACCATGGC	5781
Qy	5066	AATTCGGGCTTCTGGCAACCAAGCAATGAAACCGATGGACAAACATTTTATGAGTAC	5125
Db	5782	AATCTGGGCTTCTGGCAACCAAGCAATGAAACCGGTTGACAAACCTTTTATGAGTAT	5841
Qy	5126	GACAGCTTTGGCGGCTGACAAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGA	5185
Db	5842	GACAGTTTGGTGGCTGACAAACGTTGACCTTCCAACTGGCCAGGTGAGCAGTTTCCGA	5901
Qy	5186	AGTGATACAGACAGTTTCACTGATGTCCAGGTAGAGACTCTCAGCAAGGATGATGTCACC	5245
Db	5902	AGCGATACAGACAGTTTCACTGATGTCCAGGTAGAGACTCTCAGCAAGGATGATGTCACC	5961
Qy	5246	ATAACCAACCACTGTGCTCAGCGGCTTCTACACACTGCTGCAAGCAAGTCCCGG	5305
Db	5962	ATAACCAACCACTGTGCTCAGCGGCTTCTACACACTGCTTACAGCAAGTCCCGG	6021
Qy	5306	AACAGCTTACTACATCGGGCGGATGGCTCTCTGGGCTGTCTGGCCCAACGCGATGGAG	5365
Db	6022	AACAGCTTACTACATCGGGCGGATGGCTCTCTGGGCTGTCTAGCCCAATGSCATGGAA	6081
Qy	5366	GTGGCGTGCAGACTGAGCCCACTTGTGGCTGGCAACCGTCAACCCCACTGGGCGAAG	5425
Db	6082	GTGGCTGTGAGACTGAGCCCACTTGTGGCTGGCACTGTCAACCCCACTGTAGGCAAG	6141
Qy	5426	AGGAATGTACGCTGCCATCGCAACAGGCTCAACCTGTGTGAGTGGCGCCAGCGCAAA	5485
Db	6142	AGGAATGTACATGCCCATTGATATGCTTCACTGTGTGAGTGGCGGAGCGCAAG	6201
Qy	5486	GAGCAGGCTCGGGCCAGGTCCTGTCTTTGGCGCGGCTGCGGGTGCACAAACCGAAAT	5545
Db	6202	GAGCAGGCTCGTGGCCAGGTCACCGCTTTTGGACCCCGCTTGGCGGTTTCAACACCGAAAC	6261
Qy	5546	CTCTATCTCTGACATTTGATCGGTAAACACGACACAGAGAAGATCTTATGATGACCCCGC	5605
Db	6262	CTCTTGTCTTGGACTTTGACCGGTGTCAACGACACAGAGAAGATCTTACGATGACCATCGC	6321
Qy	5606	AAGTTCACTCTCGGATTTCTGACGACCGGCGGCGCCAGCTCTGGTCAACCCAGC	5665
Db	6322	AAGTTCACTCTCGGATTTCTTATGACGAGGAGGCGCCAGCTTCTGGTCAACCTAGC	6381
Qy	5666	AGCAGGCTGAATGGTGTCAACGTCATCTCCCTGGGGGTTTACATGCTGCGATCCAG	5725
Db	6382	AGCAGGCTGAATGGTGTAAATGTGACCTACTCCCTGGAGGTCACTTGTGGAATCCAA	6441

QY 5726 AGGGGATCATGTCTGTAAGAATGGAATACGACCGAGGGGCGGCATCACATCCAGGATC 5785
Db 6442 AGGGGATCATGTCTGAGAGAAATGGAATATGATCAGGAGGGCGGCATCACATCCGGATC 6501
QY 5786 TTGCGTGTGGAAGACATGAGGATGACATACATCTTAGAGAGTGCATGGTCTCTACTA 5845
Db 6502 TTGCGAGCGGAATAATGTGGAGCTACACGTACTTTAGAGAGTGCATGGTCTCTATCTC 6561
QY 5846 CACAGCCAGGACGATATATCTTTGAGTTGCAAGAAATGACCGCTCTCTTCTGTGACC 5905
Db 6562 CACAGCCAGGACGATATATCTTTGAGTTGCAAGAAATGACCGCTCTCTTCTGTGACC 6621
QY 5906 ATGCCCAAGCTGGCGGAGACATAGACCATCCGCTCAGTGGGCTACTACAGAAAC 5965
Db 6622 ATGCCCAAGCTAGCCCGGAGACGCTGGAGACCATCCGCTCAGTGGGCTACTACAGAAAC 6681
QY 5966 ATCTATCAGCCCTGAGGGCAATCCCTCAGTACATACAGGACTTCACTGAGGATGGGAC 6025
Db 6682 ATCTACAGCCCTGGAGAGCAACGCTCAGTCACTCAGGACTTCACTGAGGATGGACAC 6741
QY 6026 CTCCTTACACCTTCTACTCTGGGCACTGGCGGAGGTTGATACAAATATGGCAACTG 6085
Db 6742 CTGTTTACATACCTTCTACTTGGGCACTGGCGGAGGTTGATACAAATATGGCAACTG 6801
QY 6086 TCAAGAGCTGGGAGACGCTCTATGACACACCAAGGTCAGTTTCACTATGACGAGACG 6145
Db 6802 TCAAGAGCTGGGAGACGCTCTATGACACACCAAGGTCAGTTTCACTATGACGAGACG 6861
QY 6146 GAGGATCTGTAAGACCATCAACTACAGAAATGAGGCTTCACTGACCAATCCGCTAC 6205
Db 6862 GAGGATCTGTAAGACCATCAACTACAGAAATGAGGCTTCACTGACCAATCCGCTAC 6921
QY 6206 GTCAGATTGGCCCTGATTGACCGACAGATCTTCGCTTCACTGAGGAGGATGGTC 6265
Db 6922 GTCAGATTGGCCCTGATTGATGGGAGATCTTCGCTTCACTGAGGAGGATGGTC 6981
QY 6266 AACGCCCTTTGACTACAACTATGACAAACAGCTTCCGGGTGACGAGCATGCAGGCTGTG 6325
Db 6982 AATGCCCTTTGATTACAACTATGACAAACAGTTCGCTGACTAGCATGCAGGCTGTG 7041
QY 6326 ATCAAGAGACCCACTGCTCCATGATCTATGCTATGATGATGATGATGATGATGATG 6385
Db 7042 ATCAATGAGACCCACTGCTCCATGATCTATGCTATGATGATGATGATGATGATGATG 7101
QY 6386 GAGCAGTTGGGAAGTTTGGTGTCTATTACTATGACATTAACAGATCATCACACAGCT 6445
Db 7102 GAGCAGTTGGGAAGTTTGGTGTCTATTACTATGACATTAACAGATCATCACACAGCT 7161
QY 6446 GTCATGACCCACACAAAGCATTTTGTATGATGATGATGATGATGATGATGATGATG 6505
Db 7162 GTCATGACCCACTCCAAGCATTTTGTATGATGATGATGATGATGATGATGATGATG 7221
QY 6506 ATCTTCCCTCGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6565
Db 7222 ACTTTCCTGCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 7281
QY 6566 AAGAGGAGCTGAAGGTAGGACCTTACGCCAATACCACTCGCTACTCTCTATGATGATG 6625
Db 7282 AAGAGGAGCTGAAGGTAGGACCTTATGCCAATACCACTACCCGCTACTCTCTATGATG 7341
QY 6626 GCTGACGCCAGCTGCAGACAGTCTCCATCATGATGATGATGATGATGATGATGATGATG 6685
Db 7342 GCTGATGCCAGCTGCAGACAGTCTCCATCAATGATGATGATGATGATGATGATGATGATG 7401
QY 6686 GACCTCAATGGGAACCTGCACTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTA 6745
Db 7402 GACCTCAATGGGAACCTTACTGAGCCCTGGGAACAGGCAACGGCTCACACCACTA 7461
QY 6746 CGGTATGACATCCGGGACCGCATCACTCGGCTGGGTGAGTGCATACAAAGTGGATGAG 6805
Db 7462 CGGTATGACATCCGGGACCGCATCACTAGGCTGGGTGATGATGATGATGATGATGATGATG 7521
QY 6806 GATGGCTTCTGAGGACGCGGGGGTGTATATCTTTGAGTACAACTCAGCTGCGCTGCTC 6865

Db 7522 GATGGCTTCTGAGGACGCGGGGTGGGATGTCTCGAGTACAAATCAGCCGGGTGCTC 7581
QY 6866 ATCAAGGCTTACAAACCGGCTGGCAGCTGGAGTGTACAGTACCGTACGATGGCTGGG 6925
Db 7582 ATCAAGGCTTACAAACCGGCTGGGCTAGTGGAGTGTACAGTACCGTACGATGGGTGGGA 7641
QY 6926 CGCGGCTGTCCAGCAAGAGACGACACAGCCACCACTGCACTTCTTATCGACACCTG 6985
Db 7642 CGCGGATATCCAGCAAGAGACGACACAGCCACCACTGCACTTCTTATCGACACCTG 7701
QY 6986 ACCAACCCACCAAGTACCCACCTGTACACCACTCCAGCTCTGAGATCACTCCCTC 7045
Db 7702 ACCAACCCACCAAGTACCCACCTGTACACCACTCCAGCTCTGAGATCACTCCCTC 7761
QY 7046 TACTAGACTTGCAGGACACCTCTTTGCCATGGAGCTGAGCAGTGTGATGAGTTTAC 7105
Db 7762 TACTATGATCTGCAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGGATGATGATG 7821
QY 7106 ATAGCTTGTGAACAATCGGACCCCTCTTGTCTTGTAGTGAACAGGTTTGTATGATC 7165
Db 7822 ATAGCTTGTGAACAATCGGACCCCTGCTTGTCTTGTAGTGAACAGGTTTGTATGATG 7881
QY 7166 AAGCAAACTCTGACACAGCTATGGGAGATCTACATGATGATGATGATGATGATGATG 7225
Db 7882 AAGCAGATCTTATACGCGCTATGGGAGATCTACATGATGATGATGATGATGATGATG 7941
QY 7226 ATCATCATAGGCTACCATGCTGCTCTATGATGATGATGATGATGATGATGATGATG 7285
Db 7942 ATCATCATGCTACCAAGGCTCTTATGATGATGATGATGATGATGATGATGATGATG 8001
QY 7286 CGCGGAGATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7345
Db 8002 CGAGCGGATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8061
QY 7346 CACCTTAGTAGCAGCAAGCTCATGCTTAACTCTATATGTTCAAAAACCAACACCC 7405
Db 8062 CGCTGAGTAGCAACAGCATCGTCTTCTATCTATGATGATGATGATGATGATGATG 8121
QY 7406 ATCAGCACTCCAGGACATCAAGTGTCTATGACAGATGTTAAGTGTGCTGCTGCTGCT 7465
Db 8122 ATCAGCACTCTCAGGACATCAAGTGTCTATGACAGATGTTAAGTGTGCTGCTGCTGCT 8181
QY 7466 TTTGATTCAGCTACAAACGCTGATCCTGCTTATCCAAACCAAGATGATGATGATGATG 7525
Db 8182 TTTGATTCAGCTACAAACGCTGATCCTGCTTATCCAAACCAAGATGATGATGATGATG 8241
QY 7526 GAACTCTCTACGAGCTCATCCACACAGATGAAACCGCAGGAGTGGGCAACAGCAAG 7585
Db 8242 GAACTCTCTACGAGCTCATCCACACAGATGAAACCTCAGGATGGGCAACAGCAAG 8301
QY 7586 TCTATCTCGGGTACAGTGTGAGTACAGAGAGCTCAAGGCTTGTCACTGATGAA 7645
Db 8302 TCTATCTCGGGTACAGTGTGAGTACAGAGAGCTCAAGGCTTGTCACTGATGAA 8361
QY 7646 CGGTTTGAACAGCTCTATGCTGCTCACAATCAGAGTCCAGCAGGCTTCCAAAGACCAAG 7705
Db 8362 CGCTTGAACAGCTCTATGCTGCTCACAATCAGAGTCCAGCAGGCTTCCAAAGACCAAG 8421
QY 7706 AAGTTTGTATCCAGGCTCAGTCTTTGGCAAGGGGTCAAGTGTGCTTGAAGGATGCT 7765
Db 8422 AAGTTTGTATCCAGGCTCAGTCTTTGGCAAGGGGTCAAGTGTGCTTGAAGGATGCT 8481
QY 7766 CGAGTGAACACAGATCATCAGTGTGGCAATGAGATGGGCAAGGTTGCTGCTGCTC 7825
Db 8482 CGAGTGAACACAGATCATCAGTGTGGCAATGAGATGGGCAAGGTTGCTGCTGCTC 8541
QY 7826 TTGAACCAATGCCACTACTAGAACTCTGACCTTCACTTGAATGATGGGTGGATACCCAT 7885
Db 8542 TTGAACCAATGCTCACTATYTAGAACTTCACTTCACTTGAATGATGGGTGGATACCCAT 8601
QY 7886 TACTTTGTGAACACAGGACCTTCAAGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTG 7945

Db	2714	GAGCTGTGGCAGCGGACTGTGTGGGACATGGCACTGTGTGGCGGGAGAGTGCCTGGTGC	2773
Qy	2273	GAGGATGGCTGGATGGGGGACGCTGCGACACAGGGGGCTGCCACCGCGCGTGTGGCGAG	2332
Db	2774	GACGAGGCTGGATGGGCACCGGTGCGAGCAACGGGCTTGTACCCCTCGCTGCGAGGAA	2833
Qy	2333	CATGGGACCTGCGCGGACGGCAAGTGCAGTGCAGCCCTGCTGGAATGGCGAATGCG	2392
Db	2834	CACGGGACCTCAAGGACGGAAAGTGCAGTGCAGCCCGGATGGAATGGAGAACCTGC	2893
Qy	2393	ACCATCGCTCACTACTGTGGATAGGTAGTTAAAGAGGGTTCCTCGGTTGTGCAATGGC	2452
Db	2894	ACTATC-----GAGGGCTGCCCTGGTCTGTGCAATGGC	2926
Qy	2453	AACGSCAGATGACTTATAGACTGAAATGTTGGACATCGCTGCTGCGAGCTGGGTGGAGA	2512
Db	2927	AATGSCAGATGACGCTGGTAAATAATGGCTGGTACTGCGTTTGTCAACTGGGTGGCGG	2986
Qy	2513	GGAGCTGCTGTGACACATTCATGGAGACTGCGTGGGTGACAGCAAGACAAATGATGGA	2572
Db	2987	GGAGCAGGCTGGACACATCAATGGAGACGGCTGCGAGTACGCGAAGGACAAATGATGGA	3046
Qy	2573	GATGCCCTGGTGGACTGATGAGACCTGACTGCTGCTCCAGCCCCCTGTGCCATATCAAC	2632
Db	3047	GATGGGTTGACGGACTGCATGGAACCGGATTTGCTTGCAGGCATCATGTACACAACA	3106
Qy	2633	CCGCTGTGCTGGCTCCGCTAAACCTCTGACATCATCCAGGAGACACAGGTCCTGTG	2692
Db	3107	TCTCTGTGTGGGTCTCCCGATCCTCTGTGATATCATTCAGGAACCCAAATCTCTTCC	3166
Qy	2693	TCACAGCAGAACCTACACTCTCTATGACCGGCATCAAGTTCTCTGCTGGGAGGAGCAGC	2752
Db	3167	AGTCTGAGCACCTGCAGTCTCTATCAGCGCATCATTTCTTGGTGGGCGGGAGCAGC	3226
Qy	2753	AGGACATAATCCCGGGGAGAACCCCTTTGATGAGGGGATGCTGTGTATTTCTGTGGC	2812
Db	3227	ACCCATGTCACTGTGTCAATCCCTTTGATGGCATTCATGCTGCTGTATCCGGGA	3286
Qy	2813	CAAGTGATGACATCAGATGGAAACCCCTGTGTGTGGAATCAGTTTGTCAATAAC	2872
Db	3287	CAAGTGTGACTTCTGATGGAACTCTTTGGTGGGTAATATCAGCTTCATTAACAA	3346
Qy	2873	CCTCTCTTTGGATATACAAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGC	2932
Db	3347	CGGCGTATGGGTATACAAATCACAGACAGGACGGCAGCTTTGACCTGTTAAGCAACGGT	3406
Qy	2933	GGCATCTCAATCATCTCGGTTGAGGGGACCTTTTCATCACACAGGAGCACACCTTG	2992
Db	3407	GGCGTGGCCATTTGGCCTCGGTTTGGCGTCTCCCTTTATCACTCAGGAGCACACCTTG	3466
Qy	2993	TGGCTGCCATGGGATCGCTTCTTTGTCTATGGAACCAATCATCATGACATGAGAGAGAAAT	3052
Db	3467	TGGCTGGCGTGGGGCGTTTCTTGTCTATGGATACAAATTTGTATGGACACAGGTCAAC	3526
Qy	3053	GAGATCCGAGCTGTGACTGAGCAATTTTGGCGGCCCAACCCAGTGTCTCTCCATCC	3112
Db	3527	GACATCCCGCAGCTGGAGCTGAGTAGTCTACCCGTCCTCCATCGCCCATCGTCTCCCTGCT	3586
Qy	3113	CCACTGAGCTCTTCGCGAGCTCTGTGACAGAAAGCCCAATTTGTGCGGAAATTCAG	3172
Db	3587	CCTTTGACAGCGTTTGGAGGACCTGCCCTGAGAGGGGATCGTCTTCAGAGATTTCAG	3646
Qy	3173	GCTTTGAGGAGGAAATCTATATCTCTGGCTGCAAGATGAGGTGAGCTACCTGAGCAGC	3232
Db	3647	ACCCCTCCAAGAAGTGGATCCAGTACCAAGTATGCGCTTGGGCTACTTGGAGCAGC	3706
Qy	3233	CGGACCTGGCTACAAATCTGCTGAGGATCAGCTCACCCACCCAGCACCTCCCTTC	3292
Db	3707	CGGACTCTGGCTATAAGTCTCTTCTGCGAATCACCCCTCACCCACTCCACCATCCCGTTC	3766
Qy	3293	AACCTCATGAGGTGCACCTCATGTAGGTGGGCGCGCTCTTTCAGGAAGTGTTC	3352
Db	3767	AGCCTGATGAAGTGCACCTATATGTGGCGGTGGAAGAGACTTTTCCGGAAGTGTTC	3826
Qy	3353	GCTGACGCCCCAGACCTGTCTATTATTTTCAATTTGGGACAAAGACAGACGCTCTACACCCAG	3412
Db	3827	TCAGCTGCTCCCAACCTCTCTATGACTTTGTTGGGACAAAACCCAGCTCTACAGCCAG	3886
Qy	3413	AGGCTGTTTGGGCTTTTCAGAACCTTTGTTCCGTTGGGTATGAATGAATCAATCTCTCCCA	3472
Db	3887	AAGTATATATGGGCTGTCTGAGGCTTTTGTGCTCGTGGGTTTGAGTATGAATCTCTCCA	3946
Qy	3473	GATCTAATCTCTGTTGGGAAAAAGAAACAACAGTCTCTCAGGGCTATGAAATTCAGCGCTCC	3532
Db	3947	GATCTAATCTCTGTTGGGAGGAGGAGCGGCTGTTCTGACGGGATACGAGACCACTGCGTCA	4006
Qy	3533	AGCTTTGGAGATGGAGCCTAGACAAACATCATGCCCTCAAATTCAAAGTGGTATCCTG	3592
Db	4007	AATCTTTGGTGGCTGTGAGCAACCAACCATGCGCTGAACATTCAAAGTGGTATCTCT	4066
Qy	3593	CACAAAGGGAATGGGAGAACCAAGTTTGTCTCAGCAGCTCTCTGCTCATTTGGGAGCATC	3652
Db	4067	CACAAGGGCAACGGTGAGAACATCTTCATTTCCAGCAGCTCCAGTCTATAGGAAGCATC	4126
Qy	3653	ATGGGCAATGGGCGCCGGAGAGCATCTCTGCGCCAGCTGCAACCGGCTTGTGTACGGC	3712
Db	4127	ATGGGTAATGGACGGAGGGCGAGTATCTCTGTCCAGCTGCAACCGGGTGGCGGATGA	4186
Qy	3713	ACAAGCTCTGCGCCCGAGTGGCCCTCACTGTGGCTCTGACGGGAGCTCTATGTGGT	3772
Db	4187	ACAAGCTGCTGCACCTGTAGCTTTAGCTTGGGTCGATGGGAGTCTGTATGTGGA	4246
Qy	3773	GATTTCAACTACATTAAGAGGATTTCCCTCTCTGGAATGTCAACCAATCTCTAGAGCTG	3832
Db	4247	GATTTCACTACGTCGCTCGCTCTTCAACCGGAAATGTCAACAGCGCTCTAGAGCTC	4306
Qy	3833	AGGAATAAGATTTACAGATAGTACAGTCCAGCACAAATACCTACCTGSCCAGAC	3892
Db	4307	AGCA-----ACAGCCCTGACACAAATATTAACCTGGCCACAGT	4345
Qy	3893	CCCATGAGTGGGCGCTCTTCTTCTGACAGCAACAGCGCGGCTCTTTAAAAATCAAG	3952
Db	4346	CCTGTTTCAAGTTGGCTGTACTCTCTCGGACACAGCAGCAGGAAGTGTTCAGGCAAG	4405
Qy	3953	TCCACTGTGGTGTGAAGACCTTGTCAAGAACTGTAGGTGGTTTGGGGGACAGGTGAC	4012
Db	4406	TCTCTGTATGCACTAAAAGATGTGGCCAAAACCTCGAGCTGTGTGGGGAACAGAGAT	4465
Qy	4013	CAGTGCCTCCCTTTGATGACACTCGCTCGGGGATGTGGGAAGCCACAGAAAGCCACA	4072
Db	4466	CAGTGTCTGCCATATGACGAGACTCGATGTGGGATGGAGGCAAGCTGTGGAGGCCACA	4525
Qy	4073	CTCAACCAATCCAGGGGTATTAACAGTGAACAAGTTTGGGCTGATCTACTTCTGGATGGC	4132
Db	4526	CTCAACCAACCAAGAGGCACTACCGTGGATAAGTATGGTGTGATCTTCTTTGGATGGC	4585
Qy	4133	ACCATGATCAGACGATCGATCAGATGGGATCATCTCCACCTGCTCGGCTCTAATGAT	4192
Db	4586	ACCATGATCGCGGATCGATCAGAACCGGATCATATCCACCTGCTGGGCTTCAACGAC	4645
Qy	4193	CTCACATCAGCCCGGCACTCAGCTGTGATTTCTGTATGATATTTCCAGGTAAAGCTG	4252
Db	4646	TTGACTTGGCTCGACCCCTTAGCTGACTCGGTATGGACATCTCGCAGGTGCGCCTG	4705
Qy	4253	GAGTGGCCACAGACTTAGCCATCAACCAATGGACAACTCACTTTATGCTCTGCAAC	4312
Db	4706	GAGTGGCCGACTGACCTCGCTGTGAGCCCATGAGCAACTCTCTTTATGTTTGGCAAC	4765
Qy	4313	AATGTGTCTGCTCAAAATCTCTGAAAACCAACAGGTGCGCATTTGCTCGGAGGCGCCATG	4372
Db	4766	AATGTAGTCTTGAGATCTCAGNAACCATCAGTTCAGGATTTAGCAGGACGCCCATG	4825
Qy	4373	CATGCGAGGTCCCTGGCATTTGACACTTCTGCTTAAGCAAGGTGGCCATTCACGCAAC	4432
Db	4826	CATGCGAGGTCCCGGTCTGGACCACTTCTCTAGTCAGCAAAATCGCCATCCATGCCACC	4885

Dd	7046	GCAAAATACCA	CCCGCTATG	CCTATGA	TATATGTAG	CAGATGGT	CAGCTC	CAAGTAGTGTC	7105
Qy	6653	ATCAATGACA	AAGCACCTCT	CGCGCTA	CAGCTAC	GAACCTCA	ATGCGAACCT	TGCACCTACTG	6712
Dd	7106	ATCAATGATA	AAACCCTCAT	GGCGCTA	CAGTTAT	TGACCTTA	TATGCAATTT	TGACCTCCTT	7165
Qy	6713	AGCCTGGGA	CAGTGCAC	CGGCTCAC	CACACTAC	GGTATG	ACATCCGCGAC	CGCATCACT	6772
Dd	7166	AGCCTCGT	AATAGCGC	CAGACTTAC	ACCCTAC	GTACGACAT	TTCGGGAC	CGCATTTACT	7225
Qy	6773	CGGCTGGT	GACCGT	GCAATACA	GATGGAT	GAGATGGCTT	CTCTGAGG	CACAGCGGCGGT	6832
Dd	7226	CGTCTAGT	GATGTG	CGAGTAC	CGTTTTAG	ATGAGATG	GAATCTCT	CAGCAAAGAGCAAT	7285
Qy	6833	GATATCTTT	GAGTACAA	CTCAGCT	TGGCCTCG	TONTAA	GCGCTTACAA	CCGGCGTGGCAGC	6892
Dd	7286	GATTCTCTT	TTGAGTAA	CTCTGCT	TGGGTTCG	TTTAAAA	CCCTACAA	CAAAAGTTAAACGGA	7345
Qy	6893	TGGAGTGT	CTCAGGTAC	CGCTAC	GATGGCT	TGGGCG	CGCGTGTCC	ACAAGAAGACGACAC	6952
Dd	7346	TGGACTAT	CAAGTAC	CGCTATG	TGGCTT	TGGG	CAGAAGATGTCC	AGTCCAAGTATGCCAA	7405
Qy	6953	AGCAACCA	CCGTGCAG	TTCTTTAT	PGAGA	CTGTGACA	CCCAACCC	CAACAGGTCAACCCACTG	7012
Dd	7406	GGCACCAT	CTGCAAA	TTCTTCTAT	GCAGACTT	GTGCC	AGTCCCAC	CAAGTCAACACATG	7465
Qy	7013	TACAACACT	CCCGCTCT	CGAGTAC	CACTCC	CTCTACT	ACGACTT	GCAAGACACCTCTCTT	7072
Dd	7466	TACAATCA	CTCAAGCT	CAGAAAT	CACTCACT	TACTAT	GACCTTCA	AGGCCACCTTTTT	7525
Qy	7073	GCCATGAG	CTGAGCAG	TGGTATG	AGTTTTACA	TAGCTTT	GTGACAA	CTCGGGACCCCT	7132
Dd	7526	GCCATGAG	TTGAGCAG	TGGGATG	AGTTCTAT	TGTG	GCATGCG	ACAACATTGTTACTCCA	7585
Qy	7133	CTTGCTGT	CTTTAGT	TGAAACAG	GTTCAT	GTATCA	AGCAAA	TCTGTACACAGCCTATGGG	7192
Dd	7586	CTGCTCTA	TTTATAG	TGGCGT	GTGFTCTAT	GATCAAA	CAGATTC	CTCCAACAGCTTTTGGG	7645
Qy	7193	GAGATCTA	CAATGGA	TACAA	CCCCAACTTT	CAGATCAT	CTATAG	CGCTACCACTTGTGGCGTC	7252
Dd	7646	GAAGTATAT	TTAGACT	CCAACCC	NAAGCTT	CCAGCT	TAGTAATAG	CGCTACAGGCGGCGCTT	7705
Qy	7253	TATGATCCA	CTACCA	AGCTTTGTCC	ATGGG	CGCGGAGATTA	TGATGTG	CTGGCGCGGA	7312
Dd	7706	TATGAGCC	CTTTACAAA	GCCTTGCAT	TATGG	CGCGCAGAGATTA	TGATGTTCTT	TGTCGAGGC	7765
Qy	7313	CGCTGGAC	TAGCCAG	ACCAAGC	CTGTG	GAAGCACCTT	AGTAGC	ACAACGCTATCGCT	7372
Dd	7766	CGFTGGAC	CACTCTG	ACCATGATAT	CCGGA	AGCGGTCA	ACAGCG	ACAATTGTGCC	7825
Qy	7373	TTTAATCT	CTATATG	TTTCAAAA	CAACA	CCCCAT	CAGCACT	CTCCAGGACATCAAGTGC	7432
Dd	7826	TTCAACCT	CTACATG	TTTAA	GACAACA	CAACCCCTT	GAGCAAC	CGCCAGAAGAACCAAGTGC	7885
Qy	7433	TTCATGAC	AGATGTTAA	CAGCTG	CTGCTAC	CTTTGGAT	TTCAGCT	TACACAAGCTGATC	7492
Dd	7886	TACATGAC	AGATGTTAA	CAGCTG	GTGTCG	GTTCG	GTTC	CAAGCTATACAAGCTCATC	7945
Qy	7493	CCTGTTAT	CCCAAC	CAGACAT	GGAATGCC	ATGGAA	CCCTCT	CACGAGCTCATCCACACA	7552
Dd	7946	CCTGGCT	TAOC	GACGCTGT	CACAGATG	CCAT	TGAGCCTTC	GTACGAGCTCGTTCA	8005
Qy	7553	CAGATGAAA	ACGCA	AGGATGGGA	CAACAG	CAAGTCTTAT	CTCTCG	GGGTACAGTGTGAAGTA	7612
Dd	8006	CAGATAAGA	CACTCAG	AGGTGGAT	AGCA	CCAAAGT	CTGTTC	TAGGGGTGCAGTGTGAAGTC	8065
Qy	7613	CAGAAGCAG	CTCA	AGGCTTTG	TCACCTT	TAGAA	CGGTTTG	ACCAAGCTCTATGGCTCCACA	7672
Dd	8066	CAGAGG	CA	AGCTCAAG	CTTTTGTCC	GGTTGG	AGCGGTTT	TGTCAGATCTTACAGTGCAGT	8125
Qy	7673	ATCACAG	CTGCC	AGCAGGCT	CCAAG	ACCAAG	AAAGTTTG	ATCCAGCGGCTCAGTCTTTT	7732
Dd	8126	GATTCTGG	ATGCTCT	CAACCCCT	CTCTCC	ACACACTTTT	TGCAAC	CAGGAGCTTCTCTCTCTC	8185

Qy	7733	GGCAAGGGGTCAAGTTTGGCTTGAAGATGGCCGAGTGACCAACAGACATCATCATGTGTG	7792
Db	8186	GGGAAAGGTGAAAGTCGCATCCGTGAAGCGCGTTGAAGCGACATCATCAACCTG	8245
Qy	7793	GCCAAATGAGGATGGGGGAAGGGTTGTGCCATCTTGAACCATCCCACTACCTTAGAGAAC	7852
Db	8246	GCTAATGAAGATGGCCGAAGGATCGCCGAGTCCTAGACAAGGCTAGCTACCTGCGAGAC	8305
Qy	7853	CTGCACCTTACCATTGATGGGTTGGATACCCATTCTTTGTGAAACACAGACCTTCAGAA	7912
Db	8306	TTGCATTTTACCATTGACAGCTTTGGACACGCACACTCTTTGTCAAAATCAGCTCTGGTAGAG	8365
Qy	7913	GGTGAACCTGCCCATCTCTGGCCCTCAGTGGGGGGCGGCAACCTCTGAGAAATGGGTCAAC	7972
Db	8366	GGTGATCTATCTCTGTGGTATGACAGTAGGCGAGCTACGCTGGAGACAGAGTCAAC	8425
Qy	7973	GTCACTGTGTCCAGATCAACACAGTACTTAATGGCGAGGACTAGACGCTACACAGACATC	8032
Db	8426	GTAACAGTCTCGAGTCAACATGGTCTCTCGAGAGACGTTCCGCCCGCATAAACCGACATC	8485
Qy	8033	CAGTCCAGTACGGGGCACTGTGCTTGAACACACGCTACGGGACACAGTTGGATGAGGAG	8092
Db	8486	CAGATGAAATACGGTACTCTTAAGCCCTTAATGTACGCTACGGCAGCAGTGTGGATGAAGAA	8545
Qy	8093	AAGGCACGGTCTGTGAGCTGGGCCGGCGACAGAGCCGTGCGCAAGCGTGGGCCCGCGAG	8152
Db	8546	AAGTCAAGTCTGTGAGCTTGCAACACAGAGGGCGGTTGCAACCGCATGGGCCACGAG	8605
Qy	8153	CAGCAGAGCTGGGAGAGGGAGGAGGCTCGGGCCCTGGACAGAGGGGAGAGCAG	8212
Db	8606	CGCCACAGACTACGGCAAGGAGAAGAGGGCTCTCGGGCTTTGGACGACACGGGGAAGGCAG	8665
Qy	8213	CAGTGTCTAGCACAGGGCGGGTGC AAGGCTACGACGGCTTTTTCGTGATCTCTGTGCGAG	8272
Db	8666	CAGTCTCTAAGCTCTGGACGTGTACAGGGGCTACGAAGGCTTCTACATAGTATCAGTGCAC	8725
Qy	8273	CAGTACCAGAACTGTGCAGACAGCGCCAAACATCCACTTCATGACACAGACGCGAGATG	8332
Db	8726	CAGTCCACAGTTGACTGACAACATATATAAACGTCCATTTCTTGGGCACAGACTGATG	8785
Qy	8333	GGCCGGAGGTGACA	8346
Db	8786	GGACGCAGGTGACA	8799

DEBIT 7

RESULTS						PRI 28-AUG-2003
7						
HSM806812						
LOCUS	HSM806812	8993 bp	mRNA	linear		
DEFINITION	Homo sapiens mRNA; cDNA DKFP686K1107 (from clone DKFP686K1107);					
ACCESSION	BX640737					
VERSION	BX640737.1	GI:34364828				

KEYWORDS
SOURCE
ORGANISM

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 8993)
Pouskka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,
Mewes, H. W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and
Wiemann, S.

CONCEPT

CONSRITM
The German Human CDNA Consortium
Project Coordinator

TITLE
JOURNAL
Direct Submission
Submitted (26-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg GERMANY

Heidelberg, Germany
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the CDNA sequencing consortium of the German Genome Project.

This clone (DKP268kl1107) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information: <http://www.rzpd.de>

information about the clone and the sequencing project is available at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

sources

```

location="qualifiers"
1. 8993
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="11q14.1"
/clone="DKR2p68eK11107"
/clone_lib="686 (synonym
DH105; sites SfilA + Sfil
/dev_stage="fetal"

```

gene

1. 8993

1

```
/gene="DKFZp686K11107"
```

<1. .4200
/gene="DKFZp686K11107"
/note="hypothetical protein, N-terminus elongated,
differentially spliced"

```
/codon start=1
```

/product="hypothetical

/protein_id="CAE45850.1"
/db_xref="GI:34364829"
/translation="RIDQNGILISTILLGNDLITSRPILSCDSWMDISQVHLFWPTDLIAI
NPMSNVLVDNNVYLVNLSHQVRIYVAPRMHQVQFIDHFLFLSLUKVAIHATLESIA
LAVSHNGVLIYAETDEKINRIVQTTSGETSLVAGPSCDCKNDANCDFCSDGAY
ADAKTANLSLAVACADEGLVYADLGNIRIFIRIKNPFNTQWYBELSSPIQDELY
FDTQGHLYTQSLPGLDGYLYNFYTGDCGILTLIDNNGNMVARSDTGLWPLVLPD
GOVYWTGNTGNSALSKVTTQGHLEAMWYTHGNSGLLATKSNENGYWTTFYEDVSGBLT
NVTPPTGVQVSSPRSDTSSVYHVETSSKDDVYTTNLASCAFYTLIQQVQNSYYI
GADGRLUANGMEVALUQEPHLLSKGVTVGKRVNVTIDNGLLMLVWRQKQYA
RGQVYFGRILVARNHRLSLGDFDVRTEKTYIDPHKRTILLYDQAGRSLWSPSS
LKNQVNTVSPCYTAGTQRGIMSERMEYDQAGRITGRIFADGTWSTYVLEKSWLL
HSIOREYEFKDNQRLSSVMPNVAQTLETIRSVGYRNIYQEPGNAVSQIDFTE
DGLHLHTYFGLGRVIVKYGKLSLAETLVDTTKVFTYDAGMLKTNLONEGT
CTIRYQIQGLIDROI PRTEGGMVNAFPDYNONSFRVTSMQAVINETPLDILRY
DOVSGKTEQGRKFVYTDINGITITAVYTHKHFDAYGRMKVQYEIFRSLNMWTV
QDNMGRVYKGLKVPYNTATRYSYEDYADQGLQTSINDKPLRYSYDLGNLWML
SPGNSARLPTVDRIDRITRIGLQYKMDDEGLRQGGEDIFEYNSAGLLIKAYNRA
GMSVRYRVDLGRVRSKSHSHLQFVADLTNPTVHLVNHSSSETLSYLYDQ
GHLFAMELSGDEFFVACDNTGTPLAVSGTGMLKQILATYGEIYWMOTDPNFQII
GHGGLYDPLFTKLVMGRDRDIDVLAGWTSDEHLEKHLSSNNVWPNLWPKNNPI
SNSQIDKPFMTDNGWLLTFGLQHNVLPGYKPELMDAPESYEIHTQMKTEWONS
KSLYQCVQVQVFLAFVDFQGLYOSTISQQQATPKPKFSSVSGFGKQAL
KQORVYTIISVANDGRRVALLNHAHLYENLTIDGVDTYFVXPGSGDGLAIL
GLSGRRTPVYVWQVQINVLNGRRYRTDILQVGLGALCLNTRYGTTLDDEKARV
LELAPRQVLAQWAEQQLREBEGSLPAWTEGEKQVLTSTGRVQGYDGFVLSVQY
PELSDSANNHFWRQSEMR"

8953..8958
/gene="DXF2p68K11.07"
8973
genes="DXF2p68K11.07"

polvA signal

/gene="DKFZp686K11107"

polyA_site

/gene="DKFZp686K11107"

ORIGIN

Query Match	50.2%;	Score 4197.2;	DB 9;	Length 8993;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 4202;	Conservative 0;	Mismatches 8;	Indels 0;	Gaps 0;
QY	4145	CGCATCGATCAGAAATGGGATCATCTCCACCGCTCGGGCTCTAAATGATCTCACATCAGCC	4204	
Db	1	CGCATCGATCAGAAATGGGATCATCTCCACCGCTCGGGCTCTAAATGATCTCACATCAGCC	60	
QY	4205	CGGCACATCAGCTGTGATTTCTGTCAATGGATATTTCCAGGTAAGACTGGAGTGGCCACAC	4264	
Db	61	CGGCACATCAGCTGTGATTTCTGTCAATGGATATTTCCAGGTTACCTGGAGTGGCCACAC	120	
QY	4265	GACTTTAGCCATCAACCAATGGACAACATCACTTTTATGTCTCTCGACAACAATGTGGTCTGT	4324	
Db	121	GACTTTAGCCATCAACCAATGGACAACATCACTTTTATGTCTCTCGACAACAATGTGGTCTGT	180	
QY	4325	CAAAATCTCTGAAAACACACAGGTGGCGATTTGTGCGGGAGGCCCATCGACTGCCAGGTC	4384	
Db	181	CAAAATCTCTGAAAACACACAGGTGGCGATTTGTGCGGGAGGCCCATCGACTGCCAGGTC	240	

QY 5465 GTGGAGTGGCGGCGAGGCGAAGAGCAGAGCGCTCGGGCCAGGTCACTGTCTTTGGGCGCGG 5524
Db 1321 GTGGAGTGGCGGCGAGGCGAAGAGCAGAGCGCTCGGGCCAGGTCACTGTCTTTGGGCGCGG 1380
QY 5525 CTGGCGGTGCACAAACCGGAATCTCTATCTCTGGACTTTGTATCGCGGTAAACACGACAGAG 5584
Db 1381 CTGGCGGTGCACAAACCGGAATCTCTATCTCTGGACTTTGTATCGCGGTAAACACGACAGAG 1440
QY 5585 AAGATCTATGATGACCAACCGGAATCTCAACCTTCGGATCTGTACGACGCGCGGGGG 5644
Db 1441 AAGATCTATGATGACCAACCGGAATCTCAACCTTCGGATCTGTACGACGCGCGGGGG 1500
QY 5645 CCCAGCCTCTGCTCAACCGAGCAGAGGTGAATGTGTCAAGCTGACATATCTCCCTGGG 5704
Db 1501 CCCAGCCTCTGCTCAACCGAGCAGAGGTGAATGTGTCAAGCTGACATATCTCCCTGGG 1560
QY 5705 GGTTCATCTGCTGGATCCAGAGGGGCATCATGTCTGAAAGTAAGTAAGCAAGCGG 5764
Db 1561 GGTTCATCTGCTGGATCCAGAGGGGCATCATGTCTGAAAGTAAGTAAGCAAGCGG 1620
QY 5765 GCGCGCATCACATCCAGGATCTTCGTGATGGGAAGACATGGAGCTACACATCTAGAG 5824
Db 1621 GCGCGCATCACATCCAGGATCTTCGTGATGGGAAGACATGGAGCTACACATCTAGAG 1680
QY 5825 AAGTCCATGGTCTGTACTACACAGCCAGAGGCGAGTATATCTTTGATTTGACAAAGAA 5884
Db 1681 AAGTCCATGGTCTGTACTACACAGCCAGAGGCGAGTATATCTTTGATTTGACAAAGAA 1740
QY 5885 GACCGCCTCTCTCTGTGACAGATGCCCAAGCTGGCGGCGGACACACTAGAGACCATCCG 5944
Db 1741 GACCGCCTCTCTCTGTGACAGATGCCCAAGCTGGCGGCGGACACACTAGAGACCATCCG 1800
QY 5945 TCAGTGGGCTACTACAGAAACATCTATCAGCCCTCTGAGGGCAATGCCTCAGTCATACAG 6004
Db 1801 TCAGTGGGCTACTACAGAAACATCTATCAGCCCTCTGAGGGCAATGCCTCAGTCATACAG 1860
QY 6005 GACTTCATGAGATGGGCACTCTCTACACCTTCTACCTGGGCACTGGCGGAGGGTG 6064
Db 1861 GACTTCATGAGATGGGCACTCTCTACACCTTCTACCTGGGCACTGGCGGAGGGTG 1920
QY 6065 ATATCAAGTATGGCAACTGTCAAGCTGGCAGAGACGCTCTATGACACCAACCAAGTC 6124
Db 1921 ATATCAAGTATGGCAACTGTCAAGCTGGCAGAGACGCTCTATGACACCAACCAAGTC 1980
QY 6125 AGTTTCACTATGACGAGCGGCGCATGTGGAAGACCATCAACCTACAGAATGAGGCG 6184
Db 1981 AGTTTCACTATGACGAGCGGCGCATGTGGAAGACCATCAACCTACAGAATGAGGCG 2040
QY 6185 TTCACTGCACCATCCGCTACCGTCAGATTTGGGCCCCGTGATGACCGACAGATCTCCG 6244
Db 2041 TTCACTGCACCATCCGCTACCGTCAGATTTGGGCCCCGTGATGACCGACAGATCTCCG 2100
QY 6245 TTCACTGAGGAAGGCATGGTCAACGCCGTTTTGACTACAATGACAAACAGGTTCCGG 6304
Db 2101 TTCACTGAGGAAGGCATGGTCAACGCCGTTTTGACTACAATGACAAACAGGTTCCGG 2160
QY 6305 GTGACGAGATCCAGCTGTGATGAAAGAGACCCCACTGCCCATTTGATCTCTATCGCTAT 6364
Db 2161 GTGACGAGATCCAGCTGTGATGAAAGAGACCCCACTGCCCATTTGATCTCTATCGCTAT 2220
QY 6365 GATGATGTGTACGGAAGACAGAGAGGTTGGGAAGTTGGTCTATTCTATGACATT 6424
Db 2221 GATGATGTGTACGGAAGACAGAGAGGTTGGGAAGTTGGTCTATTCTATGACATT 2280
QY 6425 AACAGATCATCAACAGCTGTATGACCCACCAAGCAATTTGATGCATATGGCAGG 6484
Db 2281 AACAGATCATCAACAGCTGTATGACCCACCAAGCAATTTGATGCATATGGCAGG 2340
QY 6485 ATGAGGAGTGCAGTATGATCTTCCGCTCGCTCATGTACTGATGACACCGTCCAGTAT 6544
Db 2341 ATGAGGAGTGCAGTATGATCTTCCGCTCGCTCATGTACTGATGACACCGTCCAGTAT 2400
QY 6545 GATAACATGGGCGGAGTAGTGAAGAGGAGGTGAAGAGGTAGGACCCCTACGCCAATACCAC 6604

Db 2401 GATAACATGGGCGGAGTAGTGAAGAGGAGCTGAAGTAGGACCCCTACGCCAATATCACT 2460
QY 6605 CGTACTCTTATGAGTATGATGTGACGGCCAGCTGCAGACAGTCTCTCATCAATGACAAG 6664
Db 2461 CGTACTCTTATGAGTATGATGTGACGGCCAGCTGCAGACAGTCTCTCATCAATGACAAG 2520
QY 6665 CCACTCTGGCGGTACAGCTACGACCTCAATGGGAACCTGCACCTACTGAGCCCTGGGAAC 6724
Db 2521 CCACTCTGGCGGTACAGCTACGACCTCAATGGGAACCTGCACCTACTGAGCCCTGGGAAC 2580
QY 6725 AGTGACGGGTACACACCATACCGTATGACATCCGGACCGCATCACTCGGCTGGGTGAC 6784
Db 2581 AGTGACGGGTACACACCATACCGTATGACATCCGGACCGCATCACTCGGCTGGGTGAC 2640
QY 6785 GTGCAATACAAGATGATGAGGATGGCTTCTGAGGAGCGGGCGGTGATATCTTTGAG 6844
Db 2641 GTGCAATACAAGATGATGAGGATGGCTTCTGAGGAGCGGGCGGTGATATCTTTGAG 2700
QY 6845 TACAACCTCAGCTGGCTGTCTATCAAGGCTCAACCGGGCTGGCAGCTGGAGTGTCAAG 6904
Db 2701 TACAACCTCAGCTGGCTGTCTATCAAGGCTCAACCGGGCTGGCAGCTGGAGTGTCAAG 2760
QY 6905 TACCGTACATGAGCTGGCGGCGGTGTCCAGCAAGAGCAGCCACACCTG 6964
Db 2761 TACCGTACATGAGCTGGCGGCGGTGTCCAGCAAGAGCAGCCACACCTG 2820
QY 6965 CAGTTCTCTATGACAGCTGACCAACCCCAAGGTCAACCTGTACACCACTCC 7024
Db 2821 CAGTTCTCTATGACAGCTGACCAACCCCAAGGTCAACCTGTACACCACTCC 2880
QY 7025 AGCTCTGAGATCACCTCCCTCTACTACGATTTCAAGGACACCTCTTTGCCATGGAGCTG 7084
Db 2881 AGCTCTGAGATCACCTCCCTCTACTACGATTTCAAGGACACCTCTTTGCCATGGAGCTG 2940
QY 7085 AGCAGTGGTATGAGTTTACATAGCTTGTGACAAATCCGGACCCCTCTTGTGCTGCTTT 7144
Db 2941 AGCAGTGGTATGAGTTTACATAGCTTGTGACAAATCCGGACCCCTCTTGTGCTGCTTT 3000
QY 7145 AGTGGAAACAGTTTGTATGATCAAGCAAACTCCTGTACACAGCCTATGGGAGATCTACATG 7204
Db 3001 AGTGGAAACAGTTTGTATGATCAAGCAAACTCCTGTACACAGCCTATGGGAGATCTACATG 3060
QY 7205 GATACCAACCCCAACTTTGATGATCATAGGCTTCAATGGTGGCTCTTATGATCCACTC 7264
Db 3061 GATACCAACCCCAACTTTGATGATCATAGGCTTCAATGGTGGCTCTTATGATCCACTC 3120
QY 7265 ACCAAGCTTGTCCACATGGCGCGGAGATTATGATGTGTGCGCGGACGCTCGACTAGC 7324
Db 3121 ACCAAGCTTGTCCACATGGCGCGGAGATTATGATGTGTGCGCGGACGCTCGACTAGC 3180
QY 7325 CCAGACCAAGCTGTGGAGCACCTTAGTAGCAGCAACGTCTATGCTTTTAACTCTAT 7384
Db 3181 CCAGACCAAGCTGTGGAGCACCTTAGTAGCAGCAACGTCTATGCTTTTAACTCTAT 3240
QY 7385 ATGTTCAAAAACAACACCCCATCAGAACTCCAGGACATCAAGTGTCTCATGACAGAT 7444
Db 3241 ATGTTCAAAAACAACACCCCATCAGAACTCCAGGACATCAAGTGTCTCATGACAGAT 3300
QY 7445 GTTAAACAGCTGGCTGTCTACCTTTGGATTCCAGCTACAAACGTGATCCCTGGTTATCCC 7504
Db 3301 GTTAAACAGCTGGCTGTCTACCTTTGGATTCCAGCTACAAACGTGATCCCTGGTTATCCC 3360
QY 7505 AAACCAAGCATGATGCTATGGAAACCTCTACGAGCTCATCCACACAGATGAAACG 7564
Db 3361 AAACCAAGCATGATGCTATGGAAACCTCTACGAGCTCATCCACACAGATGAAACG 3420
QY 7565 CAGGAGTGGGACAAACAGAGTCTATCTCGGGGTACAGTGTGAAGTACAGAGAGCAGCTC 7624
Db 3421 CAGGAGTGGGACAAACAGAGTCTATCTCTCGGGGTACAGTGTGAAGTACAGAGAGCAGCTC 3480
QY 7625 AAGGCTTTCTACCTTAGAACGTTTGACAGCTCTATGGCTCCCACTCACCAGCTGC 7684

3481	Db	AAGCCTTTGTCACTTAGAACGGTTTGAACAGAGCTCTATGGCTCCACAATCAACAGCTGC	3540
7685	Qy	CAGCAGGCTCCAAAGACCAAGAAATTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTC	7744
3541	Db	CAGCAGGCTCCAAAGACCAAGAAATTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTC	3600
7745	Qy	AAGTTTTGCCCTTGAGAGTGGCCGAGTGACACACAGACATCATCAGTGTGGCCCAATGAGGAT	7804
3601	Db	AAGTTTTGCCCTTGAGAGTGGCCGAGTGACACACAGACATCATCAGTGTGGCCCAATGAGGAT	3660
7805	Qy	GGGCGAAGGGTGTGTGCATCTTTGAACCATATGCCACATACCTTAGAGAACCTTGCACCTTCAAC	7864
3661	Db	GGGCGAAGGGTGTGTGCATCTTTGAACCATATGCCACATACCTTAGAGAACCTTGCACCTTCAAC	3720
7865	Qy	ATTGATGGGTGGATACCCATTACTTTGTGAACACGAGACCTTCAGAAAGTGACCTGGCC	7924
3721	Db	ATTGATGGGTGGATACCCATTACTTTGTGAACACGAGACCTTCAGAAAGTGACCTGGCC	3780
7925	Qy	ATCTGGGCTCTCAGTGGGGGGCGCGAACCCCTTGAGAAATGGGGTCAACGTCACCTGTGTCC	7984
3781	Db	ATCTGGGCTCTCAGTGGGGGGCGCGAACCCCTTGAGAAATGGGGTCAACGTCACCTGTGTCC	3840
7985	Qy	CAGATCAACACAGTACTTAATGGCAGAGCTTAGACGCTACACAGACATCCAGCTCCAGTAC	8044
3841	Db	CAGATCAACACAGTACTTAATGGCAGAGCTTAGACGCTACACAGACATCCAGCTCCAGTAC	3900
8045	Qy	GGGSCACTGTGCTTTGAACACACGCTACGGCACCAACGTTTGCATGAGGAGAAGGCACGGGTC	8104
3901	Db	GGGSCACTGTGCTTTGAACACACGCTACGGGCACACGTTTGCATGAGGAGAAGGCACGGGTC	3960
8105	Qy	CTGAGCTGGCCCGGCAGAGAGCGCTGCGCCAAAGCTGTGGCCCGCGAGCAGCAGACTG	8164
3961	Db	CTGAGCTGGCCCGGCAGAGAGCGCTGCGCCAAAGCTGTGGCCCGCGAGCAGCAGACTG	4020
8165	Qy	CGGGAAGGAGAAAGGCTTCGCGGCTTGACACAGAGGGGGAGAAAGCAGCAGGTGCTGAGC	8224
4021	Db	CGGGAAGGAGAAAGGCTTCGCGGCTTGACACAGAGGGGGAGAAAGCAGCAGGTGCTGAGC	4080
8225	Qy	ACAGGCGGCTGCAAGGCTACAGCGCTTTTTTGTGATCTCTGTGAGCAGTACCCAGAA	8284
4081	Db	ACAGGCGGCTGCAAGGCTACAGCGCTTTTTTGTGATCTCTGTGAGCAGTACCCAGAA	4140
8285	Qy	CTGTACACAGCGCCACACATCCACTTCATGACACAGACGAGATGGGCCGAGGTGA	8344
4141	Db	CTGTACACAGCGCCACACATCCACTTCATGACACAGACGAGATGGGCCGAGGTGA	4200
8345	Qy	CAGAGGAC	8354
4201	Db	CAGAGGAC	4210

RESULT	8
AB037723	
LOCUS	8624 bp mRNA linear PRI 10-MAY-2002
DEFINITION	Homo sapiens mRNA for KIAA1302 protein, partial cds.
ACCESSION	AB037723
VERSION	AB037723.2 GI:20521827
KEYWORDS	.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Nagase, T., Kikuno, R., Ishikawa, K.I., Hirosewa, M. and Ohara, O.
TITLE	Prediction of the coding sequences of unidentified human genes. xvi. The complete sequences of 150 new cDNA clones from brain which code for large proteins in vitro
JOURNAL	DNA Res. 7 (1), 65-73 (2000)
MEDLINE	20181126
PubMed	10718198
REFERENCE	2 (bases 1 to 8624)
AUTHORS	Chara, O., Nagase, T. and Kikuno, R.
TITLE	Direct Submission

Submitted (31-JAN-2000) Osamu Ohara, Kazusa DNA Research Institute,
Laboratory of DNA Technology; 1532-3 Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL:www.kazusa.or.jp/huge/, Tel:+81-438-52-3913,
Fax:+81-438-52-3914)

On May 9, 2002 this sequence version replaced gi:7242958.
Location/Qualifiers
1..8624
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="bf00104"
/tissue_type="brain"
/clone_lib="pBCSKPW"
/note="This sequence was replaced that of fg05864 cdna as
a representative cdna sequence for KIAA1302."
1..8624
/gene="KIAA1302"
<590..3850
/gene="KIAA1302"
/note="Start codon is not identified."
/codon_start=1
/product="KIAA1302 protein"
/protein_id="BAA92540.2"
/db_xref="GI:20521828"
/tranlation="SPERLDNKPETTRYDSFGRLTNVTPFTGOVSFRSDTSSVHVQ
VFTSSXDVTITNLSASGAFVTLQDOVNSYVIGADGSLRLLANGMEVALQTEPH
LGIATVPVGRKVTLPDNLGNLVENQRQEQAVTVFGRRLLVNRNLLSLDF
DVRTREKIDYDHRKVELLKYDQAGRPSLSPSSRLNGVNVYSPGGIYGLQIGM
NMARQDQAGRTSRIFADGKWSYLYLEKSVLLHSQRYIFEGDKNDRLSSVTMP
SVARQLETRISYGVNTYQPGENASVQIDFTEDHLLHFTFLGTRGVYIKYGL
SKLAETLYRTKVSPTYSQAMGLNTLNQNEGFTCTIRYQIGLIDQIQRFTREG
MYNARDYNVDSFRTVSCAVINTEPLDILRYDDVSKGTQFQKFGVIVYDINQI
IITAYTHTKFDATGRMKVEQYEIERSLWYMTVQYDNWGVVKKELXVGVANTTR
YSIEYDAGQLOQTVSINDKPLWRYSYDLNGLNLLHLSFGNSARLPLPLRYDIRBITRLG
DVQYKMDQDFQRQGDIFENYSGALLIKYNRAGSVRYRYDGLRRVRSKSLHS
HHLPFYADLTNFTKTHLYNHSSEITSLYDLOQLHAMELSSDEFRYIACDNIQT
PLAVFSGTGLMTIKQILYATGITYMDTNPNFQIIIGYHGLGYDPLTKLVHMGRRDYD
LAGRVSPDHELWKHLSNSNVPFNLMPKNNPINSQIDKCFMDVNSWLLTFQFI
LHNVPYGPDPMDAMEPSVELIHTOMKQTQENDNSIILGVCEVQKQLKAFVTLERF
DQLYGSTITSCQAPKTKFASGVSFGVKFALKDGRVTTDILSVANEDGRVAAI
LNHARLYNDTIDGVTHYFKPQSPGSDILAIGLSGRKTFLENGVNVTSQINTV
LNGRTRRHTDIQLQYALCLNTRYGTLLDEKARVLELARQRAVRQAWAREOORLRG
EGLRAWTEGEKQVQLSTGRVQVQDGFVFVISVEQYPELSDSANNIHFMQSEMRGR"

ORIGIN

Query Match	44.1%;	Score 3680.8;	DB 9;	Length 8624;
Best Local Similarity	97.8%;	Pred. No. 0;		
Matches 3775;	Conservative	0;	Mismatches	2;
			Indels	83;
			Gaps	1;
Qy	4578	GTAAAAATGATGCCAACTGTGATTGTTTTTCTTGAGACGATGTTATGCCAAGATGCAA	4637	
Db	1	GTAAAAATGATGCCAACTGTGATTGTTTTTCTTGAGACGATGTTATGCCAAGATGCAA	60	
Qy	4638	AGTTAAATACCCCATCTTCCTGGCTGTGTGCTCTGATGGGAGCTCTACGTGGCGGACC	4697	
Db	61	AGTTAAATACCCCATCTTCCTGGCTGTGTGCTCTGATGGGAGCTCTACGTGGCGGACC	120	
Qy	4698	TTGGGAACATCCGAATTCGGTTTTATCCGGAAGAAACAAGCTTTTCCTCAACACCCAGAAACA	4757	
Db	121	TTGGGAACATCCGAATTCGGTTTTATCCGGAAGAAACAAGCTTTTCCTCAACACCCAGAAACA	180	
Qy	4758	TGTATGAGCTGTCTTCACCAATTGACACAGAGCTCTATCTGTTTGATACACCGGCAAGC	4817	
Db	181	TGTATGAGCTGTCTTCACCAATTGACACAGAGCTCTATCTGTTTGATACACCGGCAAGC	240	
Qy	4818	ACCTGTACACCCAAAGCCCTGCCACAGGAGACTACCTGTACAACTTCACCTACCTACCTGGGG	4877	
Db	241	ACCTGTACACCCAAAGCCCTGCCACAGGAGACTACCTGTACAACTTCACCTACCTACCTGGGG	300	
Qy	4878	ACGGCGACATCACTCATCAACAGACAAACAATGGCAACATGGTAAATGTCGCGCGAGACT	4937	
Db	301	ACGGCGACATCACTCATCAACAGACAAACAATGGCAACATGGTAAATGTCGCGCGAGACT	360	

QY 4938 CTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGTGTACTGGGTGACCATGG 4997
 DB 361 CTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGTGTACTGGGTGACCATGG 420
 QY 4998 GCACCAACAGTGCATCAAGAGTGTGACACACAAAGGACACAGAGTGGCCATGATGACAT 5057
 DB 421 GCACCAACAGTGCATCAAGAGTGTGACACACAAAGGACACAGAGTGGCCATGATGACAT 480
 QY 5058 ACCATGGCAATTCGGCCCTTCTGGCAACCAAAAGCAATGAAAACGGATGGACACATTTT 5117
 DB 481 ACCATGGCAATTCGGCCCTTCTGGCAACCAAAAGCAATGAAAACGGATGGACACATTTT 540
 QY 5118 ATGA----- 5121
 DB 541 ATGACAACTCAGATGGCAGGAATTAATCATCAATTTTATCATATGAAGTACAGAGAG 600
 QY 5122 -----GTACGACAGCTTTGGCGCCTTGACAAATGTGAC 5154
 DB 601 ATTAGATAACAGCCTGAAACACACAGGTACGACAGCTTTGGCGCCTTGACAAATGTGAC 660
 QY 5155 CTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCATGTCCA 5214
 DB 661 CTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTTCAGTGCATGTCCA 720
 QY 5215 GGTAGAGACCTCCAGCAAGGATGATGATCAACATACCAACCTGTGCTCAGGCGC 5274
 DB 721 GGTAGAGACCTCCAGCAAGGATGATGATCAACATACCAACCTGTGCTCAGGCGC 780
 QY 5275 CTTCTACACACTGTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGCCGATGCTC 5334
 DB 781 CTTCTACACACTGTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGCCGATGCTC 840
 QY 5335 CTTGGCGTGTGTGGCCAAACGGCATGGAGGTGGCTGCAGACTGAGCCCACTTGCT 5394
 DB 841 CTTGGCGTGTGTGGCCAAACGGCATGGAGGTGGCTGCAGACTGAGCCCACTTGCT 900
 QY 5395 GGCTGGCACCTGCAACCCACCGTGGCAAGAGGATGTACGCTGCCCATCGAACAGG 5454
 DB 901 GGCTGGCACCTGCAACCCACCGTGGCAAGAGGATGTACGCTGCCCATCGAACAGG 960
 QY 5455 CTTCAACCTGTGGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGCCAGGTCACTGTCTT 5514
 DB 961 CTTCAACCTGTGGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGCCAGGTCACTGTCTT 1020
 QY 5515 TGGGCGCGGCTGGCGGTGCAACCCGAAATCTCTGATCTGGAATTTGATCGCGTAAC 5574
 DB 1021 TGGGCGCGGCTGGCGGTGCAACCCGAAATCTCTGATCTGGAATTTGATCGCGTAAC 1080
 QY 5575 ACGCACAGAGAAGATCTATGATGACCCGCAAGTTCAACCTTCGGATTCTGTACGACCA 5634
 DB 1081 ACGCACAGAGAAGATCTATGATGACCCGCAAGTTCAACCTTCGGATTCTGTACGACCA 1140
 QY 5635 GCGGGGCGGCCAGCCTCTGGTCAACCCAGCAGGCTGAATGGTGTCAACGTGACATA 5694
 DB 1141 GCGGGGCGGCCAGCCTCTGGTCAACCCAGCAGGCTGAATGGTGTCAACGTGACATA 1200
 QY 5695 CTCCTCTGGGGTTACATTTGGTCATCCAGAGGGATCATGCTGAAAGATGGAATA 5754
 DB 1201 CTCCTCTGGGGTTACATTTGGTCATCCAGAGGGATCATGCTGAAAGATGGAATA 1260
 QY 5755 CGACAGGCGGCCAGCCTCATCATCCAGGATCTTGCTGATGGGAAGACATGAGCCTACAC 5814
 DB 1261 CGACAGGCGGCCAGCCTCATCATCCAGGATCTTGCTGATGGGAAGACATGAGCCTACAC 1320
 QY 5815 ATACTTAGAAGTCCATGGTGTGCTACTACAGCCAGGAGGATATCTTTGAGTT 5874
 DB 1321 ATACTTAGAAGTCCATGGTGTGCTACTACAGCCAGGAGGAGTATATCTTTGAGTT 1380
 QY 5875 CGACAAGATGACCGCTCTCTTGTGACAGATGCCAAACCGTGGCGGCGAGACACTAGA 5934
 DB 1381 CGACAAGATGACCGCTCTCTTGTGACAGATGCCAAACCGTGGCGGCGAGACACTAGA 1440

QY 5935 GACCATCCGCTCAGTGGCTTACTACAGAAACATCTATCAGCCCTCTGAGGCAATGCCTC 5994
 DB 1441 GACCATCCGCTCAGTGGCTTACTACAGAAACATCTATCAGCCCTCTGAGGCAATGCCTC 1500
 QY 5995 AGTCATACAGGACTTCACTGAGGATGGCACTCTTCACTTCACTTCTACTGGGCACTGG 6054
 DB 1501 AGTCATACAGGACTTCACTGAGGATGGCACTCTTCACTTCACTTCTACTGGGCACTGG 1560
 QY 6055 CCGCAGGGTGATATCAAGTATGGCAAACTGTCAAGCTGGCAGAGACGCTCTATCAGAC 6114
 DB 1561 CCGCAGGGTGATATCAAGTATGGCAAACTGTCAAGCTGGCAGAGACGCTCTATCAGAC 1620
 QY 6115 CACCAAGGTGAGTTTCACTTATGACAGCGGAGCATGCTGAAGACCATCAACCTACA 6174
 DB 1621 CACCAAGGTGAGTTTCACTTATGACAGCGGAGCATGCTGAAGACCATCAACCTACA 1680
 QY 6175 GAATGAGGCTTCACTGACCATCCGCTACCGTACGATGGCGCTGATGACCGACA 6234
 DB 1681 GAATGAGGCTTCACTGACCATCCGCTACCGTACGATGGCGCTGATGACCGACA 1740
 QY 6235 GATCTTCCGCTTCACTGAGGAAGCATGCTCAACGCCCTTTTGTACAACTATGACAA 6294
 DB 1741 GATCTTCCGCTTCACTGAGGAAGCATGCTCAACGCCCTTTTGTACAACTATGACAA 1800
 QY 6295 CAGCTTCCGCTGACAGCATGCGAGCTGTGATCAACGAGACCCCACTGCCATTTGATCT 6354
 DB 1801 CAGCTTCCGCTGACAGCATGCGAGCTGTGATCAACGAGACCCCACTGCCATTTGATCT 1860
 QY 6355 CTATCCCTATGATGATGTGTGAGGCAAGACAGAGAGTTTGGAGTGTGTCATTTA 6414
 DB 1861 CTATCCCTATGATGATGTGTGAGGCAAGACAGAGAGTTTGGAGTGTGTCATTTA 1920
 QY 6415 CTATGACATTAACAGCATCATACACAGCTGTATGATGCCACACACCAAGCATTTGATGC 6474
 DB 1921 CTATGACATTAACAGCATCATACACAGCTGTATGATGCCACACACCAAGCATTTGATGC 1980
 QY 6475 ATATGACAGATGAAGAGAGTGCAGTATGATCTTCCGCTCGCTCATGCTGATGATGAC 6534
 DB 1981 ATATGACAGATGAAGAGAGTGCAGTATGATCTTCCGCTCGCTCATGCTGATGATGAC 2040
 QY 6535 CGTCCAGTATGATAACATGGGCGGAGTGTGAGGAAGAGAGCTGAAGGTAGGACCTTACGC 6594
 DB 2041 CGTCCAGTATGATAACATGGGCGGAGTGTGAGGAAGAGAGCTGAAGGTAGGACCTTACGC 2100
 QY 6595 CAATACACTCGCTACTCTATGATGATGATGCTGACGCGCAGCTGCACAGCTCTCCAT 6654
 DB 2101 CAATACACTCGCTACTCTATGATGATGATGCTGACGCGCAGCTGCACAGCTCTCCAT 2160
 QY 6655 CAATGACAGCCACTCTGGCGCTACAGCTACGACCTCAATGGAACTCTGACTTACTGAG 6714
 DB 2161 CAATGACAGCCACTCTGGCGCTACAGCTACGACCTCAATGGAACTCTGACTTACTGAG 2220
 QY 6715 CCTGGGAACAGTGCACGCTCACACACTACGGTATGATCCCGGACCGCATCTCTG 6774
 DB 2221 CCTGGGAACAGTGCACGCTCACACACTACGGTATGATCCCGGACCGCATCTCTG 2280
 QY 6775 GCTGGGTGACGTGCAATACAAGATGATGATGATGCTTCTGAGCAGCGGCGGTGCA 6834
 DB 2281 GCTGGGTGACGTGCAATACAAGATGATGATGATGCTTCTGAGCAGCGGCGGTGCA 2340
 QY 6835 TATCTTTGAGTACAACCTCAGCTGGCTGCTCATCAAGGCTTACAACCGGGCTGGCAGCTG 6894
 DB 2341 TATCTTTGAGTACAACCTCAGCTGGCTGCTCATCAAGGCTTACAACCGGGCTGGCAGCTG 2400
 QY 6895 GAGTGTGAGTACCGCTACGATGGCTTGGGCGGCGCTGCTCCAGCAAGAGCAGCACAG 6954
 DB 2401 GAGTGTGAGTACCGCTACGATGGCTTGGGCGGCGCTGCTCCAGCAAGAGCAGCACAG 2460
 QY 6955 CCACCACTGCACTTCTTATGACAGCTGACCAACCCCAACCAAGGTCAACCACTGTA 7014
 DB 2461 CCACCACTGCACTTCTTATGACAGCTGACCAACCCCAACCAAGGTCAACCACTGTA 2520
 QY 7015 CAACCACTCAGCTGAGATCACTCCCTCTACTACGACTGTGCAAGGACACTCTTTGC 7074

Db 2521 CAACACCTCCAGCTCTGAGATCACTCCTCTACTACGACTTGCAGAGACACCTCTTTGC 2580
 QY 7075 CATGAGCTGACGACTGTGTGATGATGTTTACATAGCTTGTGACAACTCGGACCCCTCT 7134
 Db 2581 CATGAGCTGACGACTGTGTGATGATGTTTACATAGCTTGTGACAACTCGGACCCCTCT 2640
 QY 7135 TGCTGTCTTTAGTGGAAAGAGTTTGTATGATCAAGCAATCCTGTACACAGCTATGGGA 7194
 Db 2641 TGCTGTCTTTAGTGGAAAGAGTTTGTATGATCAAGCAATCCTGTACACAGCTATGGGA 2700
 QY 7195 GATCTACATGATACCAACCCCACTTTACATGATCATATAGCTTACCATGCTGCTCTTA 7254
 Db 2701 GATCTACATGATACCAACCCCACTTTACATGATCATATAGCTTACCATGCTGCTCTTA 2760
 QY 7255 TGATCCACTACCAAGCTTGTCCACATGGGCGCGGAGATATGATGTGTGCGCGAGC 7314
 Db 2761 TGATCCACTACCAAGCTTGTCCACATGGGCGCGGAGATATGATGTGTGCGCGAGC 2820
 QY 7315 CTGGACTAGCCACAGACGCTGTGGAACACCTTTAGTAGCAACACGTCATGCTCTTT 7374
 Db 2821 CTGGACTAGCCACAGACGCTGTGGAACACCTTTAGTAGCAACACGTCATGCTCTTT 2880
 QY 7375 TAATCTCTATATGTTCAAAAACAACAACCCCATCAGCAATCTCCAGGACATCAAGTCTT 7434
 Db 2881 TAATCTCTATATGTTCAAAAACAACAACCCCATCAGCAATCTCCAGGACATCAAGTCTT 2940
 QY 7435 CATGACATGTTAACAGCTGCTGCTCACCTTTGGATTCAGCTACACACGTCATCC 7494
 Db 2941 CATGACATGTTAACAGCTGCTGCTCACCTTTGGATTCAGCTACACACGTCATCC 3000
 QY 7495 TGGTTATCCCAACACAGACATGGATGCTCAACGCTTGGATTCAGCTACACACGTCATCCACACA 7554
 Db 3001 TGGTTATCCCAACACAGACATGGATGCTCAACGCTTGGATTCAGCTACACACGTCATCCACACA 3060
 QY 7555 GATGAAACGAGAGTGGGACACAGCAAGTCTATCTCGGGTACAGTGAAGTACA 7614
 Db 3061 GATGAAACGAGAGTGGGACACAGCAAGTCTATCTCGGGTACAGTGAAGTACA 3120
 QY 7615 GAAGCAGCTCAAGGCTTTGTCACTTAGAAGGCTTTGACAGCTCTATGGCTCCACAAT 7674
 Db 3121 GAAGCAGCTCAAGGCTTTGTCACTTAGAAGGCTTTGACAGCTCTATGGCTCCACAAT 3180
 QY 7675 CACAGCTGCCAGAGGCTCCAAAGCAAGATTTGCAATCAGCGGCTCAGTCTTTGG 7734
 Db 3181 CACAGCTGCCAGAGGCTCCAAAGCAAGATTTGCAATCAGCGGCTCAGTCTTTGG 3240
 QY 7735 CAAGGGGTCAAGTTTGGCTTTGAAGGATGGCGGAGTGACACACATCATCATGCTGGC 7794
 Db 3241 CAAGGGGTCAAGTTTGGCTTTGAAGGATGGCGGAGTGACACACATCATCATGCTGGC 3300
 QY 7795 CAATGAGATGGCGAAGGTTGTGCTCATCTTGAACCATGCTCCCACTACCTAGAGAACCT 7854
 Db 3301 CAATGAGATGGCGAAGGTTGTGCTCATCTTGAACCATGCTCCCACTACCTAGAGAACCT 3360
 QY 7855 GCATTCACCATGATGGGTGGATACCCATCTTGTGAACACGACCTTCAGAGG 7914
 Db 3361 GCATTCACCATGATGGGTGGATACCCATCTTGTGAACACGACCTTCAGAGG 3420
 QY 7915 TGACCTGCCATCTCTGGGCTCAGTGGGGGGCGGCAACCTTGAGAAATGGGTCAACGT 7974
 Db 3421 TGACCTGCCATCTCTGGGCTCAGTGGGGGGCGGCAACCTTGAGAAATGGGTCAACGT 3480
 QY 7975 CACTGTGTCCAGATCAACACAGTACTTAATGGGAGCTAGAGCTACACACATCCA 8034
 Db 3481 CACTGTGTGTCCAGATCAACACAGTACTTAATGGGAGCTAGAGCTACACACATCCA 3540
 QY 8035 GCTCAGTACGGGGCACTGTCTTGAACACAGCTACGGGACACGTTGGATCAGAGAA 8094
 Db 3541 GCTCAGTACGGGGCACTGTCTTGAACACAGCTACGGGACACGTTGGATCAGAGAA 3600
 QY 8095 GGCGAGGCTCTGAGCTGGCCCGGACAGAGCGCTGCGCAAGCTGGGCGCCGAGCA 8154

Db 3601 GCACGGGTCTGGAGCTGGCCCGCAGAGAGCGTGGCGCAAGCGTGGCGCGGAGCA 3660
 QY 8155 GCAGAGACTGCGGGGAAGGGAGGAGGCTGCGGGCTTGGACAGAGGGGAGAGAGCA 8214
 Db 3661 GCAGAGACTGCGGGGAAGGGAGGAGGCTGCGGGCTTGGACAGAGGGGAGAGAGCA 3720
 QY 8215 GGTGCTGAGCAGCAGCGGGGTGCAAGGCTACAGCGCTTTTTCGTGATCTCTGTGAGCA 8274
 Db 3721 GGTGCTGAGCAGCAGCGGGGTGCAAGGCTACAGCGCTTTTTCGTGATCTCTGTGAGCA 3780
 QY 8275 GTACCCAGAACTGTGACAGCGGCCCAACATCCATTCATGACAGAGCGAGATGG 8334
 Db 3781 GTACCCAGAACTGTGACAGCGGCCCAACATCCATTCATGACAGAGCGAGATGG 3840
 QY 8335 CCGGAGGTGACAGAGGAGC 8354
 Db 3841 CCGGAGGTGACAGAGGAGC 3860

AB025412 8964 bp mRNA linear ROD 08-MAY-1999
 Mus musculus mRNA for Ten-m3, complete cds.
 AB025412
 AB025412.1 GI:4760779
 Ten-m3.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (sites)
 Oohashi,T., Zhou,X., Feng,K., Richter,B., Moertgelin,M., Perez,M.T.,
 Su,W., Chiquet-Ehrismann,R., Rauch,U. and Faessler,R.
 Mouse Ten-m/Oz is a new family of dimeric type II transmembrane
 proteins expressed in many tissues
 J. Cell Biol. (1999) In press
 2 (bases 1 to 8964)
 Oohashi,T.
 Direct Submission
 Submitted (29-MAR-1999) Toshitaka Oohashi, Okayama University
 Medical School, Dept. Mol. Biol. & Biochem.; 2-5-1 Shikata-cho,
 Okayama, Okayama 700-8252, Japan
 (E-mail: oohashi@cc.okayama-u.ac.jp, Tel: +81-86-235-7128,
 Fax: +81-86-222-7768)

Location/Qualifiers
 1..8964
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Balb/c"
 /db_xref="taxon:10090"
 /tissue_type="brain"
 /dev_stage="adult"
 1..8964
 /gene="ten-m3"
 50..8197
 /gene="ten-m3"
 /codon_start=1
 /product="ten-m3"
 /protein_id="BAA77398.1"
 /db_xref="GI:4760780"
 /translation="MDVKRRPYCSLTKSRREKERYTNSADNEECRVPTQKSYSS
 ETLKAFDHYSLRYGNRVKDLVHREADYTRQGNFTLRQLGVCEATRRGVAFCAE
 MGLPHRGYSISAGSDADTENEAVMSPEHMLWGRGVKHSRSLSSRSNSALTLTDT
 BHEMRSDESEQPSNPNQPTLQPLPSHKQHPAQHPSITSLNLSITNRNOSPAP
 SVYSPFPLPTNLTSLRSFAFKSKYCSWRKTCALCAVGSVLLAILLSFTIAMHLF
 GLNHLQTTENDTFTENGKNSDTPTNTVSLPSGDNKGLGFTHTNTITSGELDIGR
 PAALAEQLQHTTPESVQLQDSWVLGNSVPLESRHFLFKTGTTTLPFTATPGYTMASG
 BHEMRSDESEQPSNPNQPTLQPLPSHKQHPAQHPSITSLNLSITNRNOSPAP
 RAIQVAPPGIIFWRSQLFDQPFQFLNLSIQKDALIGVYGRKGLPPSHQTYDFVELLD
 GSRLTARQRLNVESEARQARQSVLSEAGFIQYLDGSIWHLAFYNDGNKPEQVSN
 TIVTESVVECPNCHNGECVSGTCHCFPGFLGPDCSRACPVLCSGNGOYSKGRCLC
 FSGWKGTCDVPTTCIDPOCGGFGICIMGSCANSYKGCNEACDLDPGCSNHGV
 CTHGECHNPNWGSNCEILKTMCAQDCSGHGTYLQBSGCTCDPNWTGPDSCSNEICS
 VDCSHGVCMGSCRCSEEGWTGPACNRACHPRCAEHGTCKDGKCECSQGWNGHCTII

FEATURES
 source
 1..8964
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Balb/c"
 /db_xref="taxon:10090"
 /tissue_type="brain"
 /dev_stage="adult"
 1..8964
 /gene="ten-m3"
 50..8197
 /gene="ten-m3"
 /codon_start=1
 /product="ten-m3"
 /protein_id="BAA77398.1"
 /db_xref="GI:4760780"

gene
 CDS

AHYLDKILKEGCGELNSNGRCTLQDNGHCHVCQPGWGRGAGCDVAMETLCTDSCDNKEJ
 DGLIDMDPCLLQSCNQPYRCGLQPDQITISGSLTQPCQAASCFYDRI3PLTJGS
 DZHLVPGESFNKSLASVIRQVLTARQDQDTPLIGVNFVSGYVITTRQCDMDL
 VANGSLATLVFRSSPLFQYHTYIMPNVFWYMDTLWKKENEDIESCDLSGVFRPS
 PIVSSPLSTVFRSSPEDBSI1PETQVILHEEYITIGDTLKLXLSISRAAGSKV/LKLT
 MTQAVIPNLMKWMLAAVUGBLFORFWPASPNSPLATYITDOKTDAYNKVYGVISEAV
 SUGVYESCLDILTLWEKRTAVIGYELDASNGWCTDLKHVILUQNGILYKNGENQ
 FISOQPPVSSIMNGRRKSI3CPSGNGQKULLAPVALACGIDSLGVYQFVNFYR
 RIFPSNGVYBELNRDHPHNSPAHRYLIVATPVDGLYVSDYNTNTRIYRPSKSI
 AKOITLKNAAVYACGTQBCQIFDEAGXGCGKAVEATLMSPKGMAIDKNGILYFVDTGM
 IRKVDQNGIITLISLNGDITSRPCTDTSMTIISOVRLEWTEPDLAINPMSSIVYLDN
 NVLQJITENRQYIAAGRPMMHQPEVEYPSGHAKVQITLESATIAVSYGVLYITE
 DEDKILNRUQVITDGEISLAVIGIPSECDKNANDCVQSGDSQAKDANAPSSLA
 ASDPTKYLADLGNBIR1RAVSNKPLISMSEYFVNDVSDOELYIPDINGHOY7WSL
 YTCGDLVYNSYNDNDYVATDUSNGTILIRDPNRMVFRVSPDQVILWITITGNGE
 LKMTAQGLVILVITFHGNSGLLAKDSBETFTFVFDVDSGRUVTITPFGVWNLH
 GMDKAT1VDEISSREBDSI2SNLSISDFYTFPDVDSGRUVTITPFGVWNLH
 VGNDSYQVBEHLVLAGTANTYKDRNLTIPGNGQNLWEFRKEQAQKVNYPGFKRL
 WLOSRLVYEDRTTKEKITYAHRKFLRLIARZITDGHPTLWLESKLMANVYVAST
 QGATSI3QRTGTSKBYDSDQGI3SRVFAADGWTSYVLEKSWMLHLQSORQI1FEVD
 MWDRSLATMPSVARHTMOT1RSIGYRNIYNDPESNAI1ITDYNESGILLOIATLFGT
 SRVLV1KYRQ2PBLSLVJSDTSVFTVDETAGUKTVNLOSOGFICTIRYQIGBELI
 DRQIFREDGCMNAFDFSYNSFRVMSQVINEPMLPILITQPDQIDSGKVRQFK
 FGVIYIDQILSTAVNTYKHFDAHRIKEIYVIEFRSLMWITIQDNGKVRQFK
 IKTPGANT1KAYEYVDQGLQVILNKEIMRWYVNDLNGHLNPPSSARITLPLR
 YUDRARI1RGDYXEDBELRGQRTIEIYEXSGKLLTVYSGKSSAVIT1VRYDGL
 GRVYSSTIGLQHLQFVADMLTP1R1THVYNHSSSETSI3YVLOQHLFAMEIISGD
 EYV1ASDNTQPLFVSSNGLMKOTQYATGEVIFSDNSVDQFOLVIFHGSLDPLTK
 LIHFGRDYDILAGRWTPD1SIWKRI1GKDPAPNLVFMNPNNSKILHOKVQYITDV
 NSLWYVGHPLHNA1IGPFVXPEDITLTSYBELVKSQWEDVP1FGVQOQVAKAKF
 LSLGKMAEQVSRKKAQAQSCMW1FAKSLIGKGVMLVYSGQVATVNLVIANEDCI
 KVAALVNNAPY1ENLHFT1EGSDTHYF1K1TTPSDSDGLTURLISGRAJENNVQ3
 QSTAVTNGRTRFADYEMQF6ALALHYRMTGLDEKARLIEQARQALANARW3Q5
 VRVDGBEGARLWTEGSKROLLSAGKGVQGVQDGYVYLSEBQVPELADSNANN1QFURQSEI
 GKR1

ORIGIN

```
Query Match      42.4%; Score 3544.6; DB 10; Length 8964;
Best Local Similarity 65.6%; Pred. No. 0;
Matches 5458; Conservative 0; Mismatches 2679; Indels 186; Gaps 12;
```

35	ATGACGCTGAAGAGAGAGAAAGCCTTACCGCTCGCTGACCCGGCGC---CGCGACGCCGAG	91
Db		
50	ATGATGTGAAGGAACGAGGCCCTTACTGCTCTCTTGACCAAGAGCAGACGGGAAGAAGAA	109
92	CGCCGCTACACAGCTCGTCCGCGGACAGCGAGAGGGAAGGCC---CGCAGAAATCG	148
Db		
110	AGCGCTATACAAATTCGTCCGCGGACAAATGAGGAGTGTAGGCTCCCCACGCGAGATCC	169
149	TACAGCTCCAGCGAGACCTGAAGGCCTACGACGAG---ACGCCGCGCTAGCCTATGGC	205
Db		
170	TATAGTTCAGTGAACCTTGAAGAGCTTCGATCATGATTATCACGCGCTCTTTATTGGA	229
206	AGCCGCGTCAAGGACATTGTGCCGAGAGGCGGAGGAATTCGCCGACAGAGTGCACAAC	265
Db		
230	AACAGAGTAAAGGATTTGGTCCACAGAGAAGCCGACGAGTATACTAGACAGGACAGAA	289
266	TTACCCCTCGGGAGCTGGGGCTCGAAGAGTAGCGCCCCCTCACGGGACCCCTGACGG	325
Db		
290	TTTACCTTAGGCAATTAGNGTGTGNAATCCGAACTCGAAGAGAGTGGCAATTCGT	349
326	ACAGACATTTGGCTGCCCAATGCGGTACTCCATNGGGGGTGCTCTGATCGCGACATG	385
Db		
350	GCAGAAATGGGGCTCCCTCACAGAGGTTACTCCATCAGTCAGGGGTGAGATCGGGATACG	409
386	GAGGCTGACACGGTGTGTCCTCCCTGAGCACCCCGTGGCTCTGTGGGGCGGAGCACACGG	445
Db		
410	GAAAAAGAGCAGTGATGTCCCCTTGACATGCCATGAGACTTTTGGGCGAGGGGGGTCAA	469
446	TCAGGGCGCAGCTCCTGCCTGTCCAGCGGGCGCAATTCCAAATCTCACTCACCGCAACC	505
Db		
470	TCGGGCGCGAGTTCCCTGCCTGTCAAGCGCGGTCCAACTCCGCGCCCTCAACCTTGACAGACG	529

[illegible]

2537	ATCATTAGCCAAAGCCCTTACAGACACCAATCTCAGAACGCTCCAAAGTCCCTCTATGACCCGA	2599
2726	ATCAAGTTCTCGTGGCAGGCAGACGACACATAAATCCCCGGGAGAAACCCCTTTGAT	2785
2597	ATCAGTTTCTGATTGGATCGGATAGCAACCAACGTCCTCCCTGGAGAAAGTCGCTTCAAT	2656
2786	GGAGGGCATGCTTGTGTATTCTGTGCCAAGTGATGACATCAGATGGAAACCCCCCTGGTT	2845
2657	AAGAGTCTTCGCTCGGTCATCAGAGGCCAAGTACTAACAGCTGATGGAAACCCCACTTATT	2716
2846	GGTGTGAACATCAGTTTGTCAATAACCCCTCTCTTTGGATATACAATCAGCAGGCAGAGT	2905
2717	GGCGTCAACGTGTCTGTTTTACACTACTCGGAATATGGAATATACCAATACCCGCCAGGAT	2776
2906	GGCAGCTTTGACTTGTGTGACAAATGGCGGCAATCTCAATCATCTCGGGTTGAGCGGGCA	2965
2777	GGAAATGTTGACTTGTGTGGCAATGTGTGGCTTCTCTGACTTTGTGTATTGTGACGCTTC	2836
2966	CCTTTCATCACAGGAGCACACCTGTGCTGCATGGGATCGCTCTTTGTTCATGGAA	3025
2837	CCATTCTCTACTAGTACCAACATGTGTGAAATCCCTGGAAATGCTCTTTATGTCATGAT	2896
3026	ACCATCATATGACATGAGAGGAAATGAGATCCCAAGCTGTGACCTGAGCAATTTTGCC	3085
2897	ACCCTTGTCATGAAGAAAGAGGAAACGACATCCCAAGCTGTGACCTCAGTGGCTTTGTG	2956
3086	CGCCCCAACCCAGTCGCTCTCCATCCCCACTGACGTCCTTCCGCAAGCTCCTGTGCAGAG	3145
2957	AGGCCAAGTCCCATCATGTGTCTTACCGTTATCCACCTTCTTCAGGTCCTTCCCCTGAG	3016
3146	AAAGCCCCCATTTGTGCCGGAAATTCAGGCTTTGCAGGAGGAAATCTCTATCTCTGGCTGC	3205
3017	GACAGCCCCATCATCCCCGAGACACAGCTCCTGCATGAAGAAACCAAAATCCAGGAACA	3076
3206	AAGATGAGGTGAGCTACCTTGACGACGCGGACCCCTGGCTACAAATCTGTCTCCTGAGATC	3265
3077	GAATTGAAACTTTCTTACCTAGTTCCTCAGAGCGGACGGGTACAAGTCAGTTCCTTAAGATT	3136
3266	AGCCTCACCCACCGACCATCCCTTCAAACCTCATGAAGGTGCAACCTCATGTGTAGCGGTG	3325
3137	ACCATGACCAAGCCGTCATACCGTTTAACTCATGAAGGTCCATCATGTGTCGCGCTG	3196
3326	GAGGGCCGCTTTACGGAAGTGGTTGCTGTAGCGCCCAAGACCTGTCCCTATTATTTCATT	3385
3197	GTTGGGAGACTCTTCCAGAAAGTGGTTTCTGCTCGCAAACTTTGGCCTACACGTTCACT	3256
3386	TGGCACAGACAGACGTCACAACGAGAGTGTTTGGGCTTTTCAGAAAGCCTTTGTTTCC	3445
3257	TGGGATGAAGCGGACGCAATATATCAGAAGTCTACGGCTTGTACAGGCGAGTTGTCTCC	3316
3446	GTGGGTTATGAATATGAATCTGCCAGATCTAATCTGTGGGAAAAAGAACCAACAGTG	3505
3317	GTCCGATACAGTACGAGTCGTGTTGGACCTGACTCTCTCGGAAAAAGAGGACTCCGCTT	3376
3506	CTGCAGGCTATGAATATGACGTCGTCGTCGAGCTTCGAGGATGGAGCCTAGACAAACATCAT	3565
3377	TTGCAAGGCATGAGTTGGATGCTTTCGAAATGTCGCGCTGGACGTTGGACACAGCACCAT	3436
3566	GCCCTCAACATTTCAAAGTGGTATCCTGCACAAAGGGAATGGGAGAACCAAGTTTGTGTCT	3625
3437	GTACTGGACGTTTCAAGCCGTTATACTATACAAGGAAATGGAGAAATCAGTTTCATCTCT	3496
3626	CAGCAGCTCTGTTCATTGGGAGCATCATGGGCAATGGGCGCCGAGAGAGCATCTCTGCG	3685
3497	CAGCAGCCTCCGGTGGTCAGCAGCATCATGGGTAAATGGTCGGAGCGGTAGCATCTCATGC	3556
3686	CCCAGCTGCACCGCCCTTGCTGACGGGCAACAAGCTCTGCGCCCCAGCTGGCCCTCACCTGT	3745
3557	CCAAGTTGCAATGGTCAAGCTGACGGGAACAAACTCTCTGGCACCCGTCGCGCTTGCCTGT	3616
3746	GGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAACTACATTAGAAGATCTCCCTCTCT	3805
3617	GGGATCGACGCGCAGTCTATACGTAGGGGATTTCAATACGTCCCGCGGAGATCCCGCTCT	3676

QY 3806 GGAAATGTCACCAACATCTCTAGAGCTGAGGAATTAAGATTTTCAGACATAGTCACAGTCCA 3865
 Db 3677 GGAATGTGACAGTGTGTTTAGAATTAAGATTAAGATTTTAGACATAGTAGAACCCA 3736
 QY 3866 GCACCAAAATCTACTCTGGCCACAGACCCCAATGAGTGGGCGCTTCTCTTTCTGACAGC 3925
 Db 3737 GCTCAGAGATCTACTGGCTACGACCCAGTCAACCGGAGATTTGTACGCTCTGTACT 3796
 QY 3926 AACACCGCGGGTCTTTAAATCAAGTCCACTGCTGGTGGTGAAGGACCTTGTCAAGAC 3985
 Db 3797 AACACCGCGAGATCTATCGGCGGAAATCACTCAGGAGGCAAGACTGACTAAAC 3856
 QY 3986 TCTGAGGTGGTTCGGGGACAGGTGACCAAGTCCCTCCCTTTGATGACACTCGTGGGG 4045
 Db 3857 GCTGAAGTGGTGGCAGGACCGGGGAAACAGTGGCTTCCCTTTGACGAGGCCAGTGTGG 3916
 QY 4046 GATGTGGGAAGGCCACAGAGCCACTCACCATCCAGGGGTATTTACAGTGGACAAG 4105
 Db 3917 GATGAGGCAAGGCTGTGGAAGCAACGCTCATGATGCCAAGGAATGGCAATGCATAAG 3976
 QY 4106 TTTGGGCTGATCTACTCTGCTGATGGCCACCACATGATCAGACGATCGATCAGAAATGGGATC 4165
 Db 3977 AACGACTGATCTACTTTGATGGAACCATGATCAGAAAGTTGATCAAAATGGATC 4036
 QY 4166 ATCTCCACCTGCTCGCTCTAATGATCTCATCAGCCCGGCCACTCAGCTGATCT 4225
 Db 4037 ATATCAACTCTCTGGGCTCCAAACGACCTCAGTCACTGACCTTTAACTGTGATCT 4096
 QY 4226 GTCATGGATATTTCCAGGTAGACTGGAGTGGCCACAGACTTAGCCATCAACCCATG 4285
 Db 4097 AGCATGCAATCAGCCAGGTGGTCTGGAAATGGCCACTGACCTCGGATCAACCCCATG 4156
 QY 4286 GACAACTCACTTTATGCTCTCGACAAATGATGCTCTGCAAAATCTCTGAAACCAACAG 4345
 Db 4157 GATAACTCCATCTACGTCCTGATAATTAAGTAGTTTTFACAGATCACTGAAACCGTCAG 4216
 QY 4346 GTGGCAATGTGCGCGAGGCCCATGCACTGCCAGTCCCTGGCATGACCACTTCGT 4405
 Db 4217 GTCGCACTGCTGCGCGCGGCCCATGCACTGTCACTGCTGAGTGGT---AATACCG 4273
 QY 4406 TTAAGCAAGTGGCCATCCACGCAACCTGAGTCAGACCGCTTTGGCTGTTCACAC 4465
 Db 4274 GTGGGAGACGCGGTTACAGCCACCTTGAGTCAGCCAGGCCATGCTGTCTAC 4333
 QY 4466 AATGGGTCCTGTATATTTGCTGAGACTGATGAGAAAAGATCAACCGCATCAGGAGTC 4525
 Db 4334 AGCGGGTCTTTATCATCAGGAAATCTGATGAGAAAGATCAACCGAATAAGCAGTC 4393
 QY 4526 ACCACTAGTGGAGATCTCACTGCTGCTGGGCCCCAGTGGCTGTGACTGTAAAT 4585
 Db 4394 ACGACAGCGGGAGATCTCTTAGTGGCTGGATACCTTCGGAATGTGCTGCAAGAC 4453
 QY 4586 GATGCCAATGTGATTTTCTGGAGACGATGTTATGCAAGGATGCAAAAGTTAAAT 4645
 Db 4454 GACGCCAATGTGACTGTACCAAGCGGAGACGCTACGCCAAGATGCCAACTCAAT 4513
 QY 4646 ACCCATCTCTGCTGCTGTGCTGATGGGAGCTCTACGTGGCGGACCTTTGGGAAC 4705
 Db 4514 GCGCGTCTCTGCTGCGCTCGCAGATGGCACTGTATCATTTGAGATCTCGGAAAT 4573
 QY 4706 ATCCGAATCGGTTTATCCGGAAGAAACAAAGCTTCTCTCAACCCAGAACATGTATGAG 4765
 Db 4574 ATCAGATCCGGCGGTTTCGAAAGATTAACCTTTACTGAACTCAATGAATTTAGAA 4633
 QY 4766 CTGCTTCCAAATTTGACAGAGCTCTATCTGTTTGTATACACCGGCAAGCACTGTAC 4825
 Db 4634 GTTGCTCTCCAACTGATCAAGAGCTCTACATCTTTGACATCAACCGTACTCACAGTAC 4693
 QY 4826 ACCCAAGCCTGCCACAGGAGACTTACCTGTACACTTCACTTACCTGGGAGCGGAC 4885
 Db 4694 ACCGTGACCTGGTCACGGGTGACTTACTATATAATTTTAGTACAGCAATGACATGAC 4753

QY 4886 ATCAGACTCATCAGACCAACATGGAATGTTAAATGTTCCGCGAGACTCTACTGGG 4945
 Db 4754 GTCACCGCTGTAATGACAGCAATGGAACACCTCCGAATCGAAGGATCGAATCGG 4813
 QY 4946 ATCCCGCTCTGGCTGGTGTCCAGATGGCCAGGTGTAAGTGTGAGTACCTGAGCAAC 5005
 Db 4814 ATCCCGTGGGGTGGTGTCTCTGATTAACAGGTGATGTTGATGATGATGATGATGATG 4873
 QY 5006 AGTGCACTCAAGAGTGTGACCAACAGGACAGGAGTGTGGCCATGTGATACCAATGCC 5065
 Db 4874 GGTGTCTGAAAGCATGACCGCTCAGGCGCTGGAATCTGGTTTGTATTACTTACCATGCC 4933
 QY 5066 AATTCGGCGCTTCTGGCAACCAACAAATGAATAAGGATGGAACATTTTATGATGATC 5125
 Db 4934 AACAGTGGCTTTTACCGCAACCAAGTGAAGAACTGGATGGAACATTTTGTACTAT 4993
 QY 5126 GACAGCTTTGGCGCTGTACAAATGATGACCTTCCCTACTGGCCAGGTGAGCACTTCCGA 5185
 Db 4994 GACAGTGAAGGTGCGCTGAGAAATGTTACTTCCCACTGGGGTGGTTTCAAACTGCCAC 5053
 QY 5186 AGTGATACAGACAGTTCAGTGATGTCAGGTAGAGACCTCCAGCA---AGGATGATGTC 5242
 Db 5054 GGGGACATGGACAAGCTATCAGGTGGACATCGAGTCATCCAGCAGAGAGGAGATGTC 5113
 QY 5243 ACCATTAACCAACCTGCTGCTCAGGGCTTCTACACTGCTGCAAGCAACGATC 5302
 Db 5114 AGCATCACTTCGAATGCTCTCCATCGATTCCTTACACCATGCTCCAGACCACTTA 5173
 QY 5303 CGGAACAGCTTACTACATCGGGCGGATGCTCTGCTGCGCTGCTGCGCAACGCGATG 5362
 Db 5174 AGAAACAGTTACAGATTTGGGTATGATGCTCTCTAGAACTTTCTATGCCAGTGTCTG 5233
 QY 5363 GAGGTGCGCTGAGACTGAGCCCACTTGTGGTGGCACCTCAACCCCACTGGGCG 5422
 Db 5234 GACTCTCACTACAGACAGACCGCCACTTCTGGTGGCACGCGAATCCACAGTAGCC 5293
 QY 5423 AAGAGAAATGTCAGCTGCGCCATCAGCAAGGCTCAACTGCTGAGTGGCGCCAGCGC 5482
 Db 5294 AAAAGAAACATGACTCTTCCCGTGAGAACGCGGCAATCTGCTGAGTGGAGTTCCGA 5353
 QY 5483 AAAGACAGCTCGGGCGGCGCTGCTTCTGGCGCGGCTGCGGCTGCACACCGA 5542
 Db 5354 AAGAAACAGCCAGGCAAGATCAACGTTATTCGGCCGGAAGCTCAGGGTCAATGGGCGC 5413
 QY 5543 AATCTCTATCTGAGCTTTGATGCGGTAAACAGCAGACAGAGAAATCTATGATGACCA 5602
 Db 5414 AACCTACTCTCAGTGACTTTGATCGGACCAACAGACGGAAGATCTATGATGACCA 5473
 QY 5603 CGCAAGTTCACCTTCGAGTTCGTACGACCGGCGGCGGCGGCGCTCTGCTCAGCC 5662
 Db 5474 CGGAATTTCTCTGAGGATCGCTTACGACAGTGGGGCAACCGACTCTTGGCTGCGG 5533
 QY 5663 AGCAGAGCTGAATGCTCAACGTAACATCTCCCTGGGGGTTTACATTTGCTGGCATC 5722
 Db 5534 AGTAGCAAGTAAATGCGAGTGAACGTCACCTACTCATCCCGTCAAATTTGCCAGCATC 5593
 QY 5723 CAGAGGGGATGATGCTGAAAGATGGAATAGCAGCAGCGGCGGCGCATCATCCAG 5782
 Db 5594 CAGAGGGGACCAACGAGCGAAAGGTGGACTATGACAGCCAGGGAGGATCGTATCTCGG 5653
 QY 5783 ATCTTCGCTGATGGGAAGACATGAGCTACATATTTAGAAAGTTCATTTGCTGTGCTA 5842
 Db 5654 GTCTTTGCCGATGGGAACATGAGTTACAGTACTTTGAAAGTCCATGTTCTTCTG 5713
 QY 5843 CTACAGCAGCAGGAGGATATATCTTTGAGTTGCAAGAAATGACCGCTCTCTTCTGTG 5902
 Db 5714 CTCCATAGCCAGCGGAGTACATCTTCGAATACGACATGTGGAGCGGCTGTCCGCAATC 5773
 QY 5903 ACCATCCCAACGTGGCGGAGACACTAGAGACCATCGCTCAGTGGGCTACTACAGA 5962
 Db 5774 ACCATGCCAGTGTGCTGCCACACCATGCGAGCAATCCGGTCCATTCGCTACTACCG 5833
 QY 5963 AACATCTATAGCCCGCTGAGGCAATGCTCAGTCAATACAGACTTCACTGAGGATGGG 6022

5834 AACATCTACATCCCGAAGAAAGCAATGCCCTCTATCATCCAGCTACACAGAGAGGG 5893
 6023 CACCTCCCTTACACCTTTACTCTGGGCACTCGCGCAGGGTGTATATCAAGATATGGCAAA 6082
 5894 CTGCTTCTGCAACACAGCTTTCTGGAACGAGTCGGAGGGTCTTATTCAAGATATAGAAG 5953
 6083 CTGTCAAGAGCTGGCAGAGACGCTCTATCACACCCACCAAGGTCAAGTTTCACTATGACGAG 6142
 5954 CAGACCGAGCTATCAGAAATTTATACACACAGCAACAGAGTCAGTTTACCTACGACGAA 6013
 6143 ACGGAGGAGCTGTGAAGACCATCAACCTACAGAAATGAGGCTTCACTGACCATCCG 6202
 6014 ACAGCGGAGTCTGAAACAGTAACCTTCAGAGTGATGTTTTATTGACCATAGA 6073
 6203 TACCGTCAGATGGCGCCCTGATTGACCGACAGATCTTCGCTTCACTGAGGAAGGATG 6262
 6074 TACAGGCAATTTGGTCCCTGATTGACAGACAGATTTTCCGCTTCAGCGAGATGGAATG 6133
 6263 GTCAACGCGCTTTTGACTACAACTATACCAACAGCTTCCGGGTACACAGCATCGAGGCT 6322
 6134 GTAAATGCGAGATTGACTATAGCTACGACACAGCTTTCGAGTGACAGCATCGAGGT 6193
 6323 GTGATCAAGGAGACCCACTGCGCATTTGATCTCTATCGCTATGATGATGTCTAGGCAAG 6382
 6194 GTCATCAATGAACACCACTGCCATTTGATCTATACCAAGTTTGTATGATCTCTGCGAAA 6253
 6383 ACAGAGCAGTTTGGGAAGTTTGGTGTCTATTACTATGATTAACCAAGATCATCAACACA 6442
 6254 GTGAGCAGTTTGGAAATTTGGAGTGATATACAGCATCAACCAAACTATTTCCAG 6313
 6443 GCTGTCAATACCAACCAAGCAATTTGATGATATGCGAGGATGAAGAGTGCATAT 6502
 6314 GCGGTGATGACTTATACAAAGCACTTTGATGCTCATGGGCGCATCAAGGAGATCCAATAT 6373
 6503 GAGATCTCCGCTCGCTCATGTGATGATGACCGTCCAGTATGATATACATGGGCGGAGTA 6562
 6374 GAGATATTAGTGTCACTCATGTACTGTGATTAACAATTCAAATATGATTAATATGGGCGGGTA 6433
 6563 GTGAAGAGAGGAGCTGAAGTAGGACCTACGCGCAATACCACTCGCTACTCTCTATGATAT 6622
 6434 ACCAAGAGAGATTAATAATTTGGGCTTTTGCCAACTACCAAAATACGCGTAGAGTAC 6493
 6623 GATGCTGACCGCAGCTCGACAGCTCTCACTCAATGACAGACCACTCTGCGCTCAGC 6582
 6494 GACGTCGATGGACAGCTCCAAACAGTTTACCTAAACGAAAGATCATGTGCGGCTACAAC 6553
 6683 TACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGGGAACAGTGCACGCTCACACCA 6742
 6554 TACGACCTTAATGGAAACCTCCACTTGTCAACCCAGCAGCGCGCCGCTGACCCCT 6613
 6743 CTACGGTATGATCCGCGACCGCATCTCGGTGGGTGAGCTGCAATACAGATGAT 6802
 6614 CTGCGCTATGACTCGGACAGAAATCACCGGCTGGGCGATGTTCAGTACCGGCTGGAT 6673
 6803 GAGGATGCTTCTGAGCAGCGGGCGGTGATATCTTTGAGTCAACTCAGCTGGGCTG 6862
 6674 GAAGATGTTTCTGCTGAGAGGGGCACTGAAATTTTGAATACAGCTCAAAAGGGCTT 6733
 6863 CTCAACAGGCTACAAACCGGCTGGCAGCTGGAGTGTGAGTACCGCTACGATGGGCTG 6922
 6734 CTGACTCGAGTCTACAGTAAGGCGAGTGGCTGGACAGTGTATCTATCGGTACGACGGGCTG 6793
 6923 GGGCGGCGCTGCTCAGAGAGAGCCACAGCCACCTGCGAGTCTTCTATGACAGAC 6982
 6794 GAAGACGTTTCTTAGCAAAACCGAGCTGGACAGCACTTCACTTTTCTACGCGGAC 6853
 6983 CTGACCAACCCCAAGGTCACCAACCTGTACAACTCAGCTCTGAGATCACCTCC 7042
 6854 CTGACATACCCCAAGGATTAATCACTGCTTACAACTTCCAGTTCAAGAAATCACCTCC 6913
 7043 CTCTACTAGACTTCCAGAGGACACTCTTTGCGCATGGAGCTGAGGAGTGTGATGTTT 7102

6914 CTGTACTATGACCTCCAGGACATCTCTGCGCATGAGATCAGCAGTGGGATGAGTTC 6973
 7103 TACATAGCTTTGACAAACATCGGGACCCCTCTTGTGCTGCTTTAGTGGAAACAGTTTGTATG 7162
 6974 TACATCGCTTCGGACAAACACGGGGACACCGCTGGGTGTTTTTTCAGCAGCAACGGGCTCATG 7033
 7163 ATCAAGCAATCTCTGTACACAGCTATGGGAGATCTACATGGATACCAACCCCAACTTT 7222
 7034 CTGAACAGACCCAGTACACTGCTTATGGTGAGATCTACTTTGACTCCAAACGCTGACTTT 7093
 7223 CAGATCATATAGGCTACCATGGTGGCTCTATGATCCACTCACCAAGCTTTGTCCACATG 7282
 7094 CAGCTGGTAATTTGATTTCCACGGGGCTTGTATGACCCGCTCACCAACTAATCCACTTT 7153
 7283 GCGCGGAGATTTATGATGTGCTGCGGACCTCGGACTAGCCAGACACAGAGTGTGG 7342
 7154 GGAGAAAGAGATTATGACATTTTGGGGAGATGGAACACACCGGACATTTGAATCTGG 7213
 7343 AAGCACTTTAGTAGCAGCAACCTCATGCTTTTAAATCTCTATATTTCAAAACCAACAC 7402
 7214 A---AAGGATCGAAAGGACCTGCTCTCTTTTAACTGTATATTTTCGGAATAACAC 7270
 7403 CCATCAGCAACTCCAGGACATCAAGTCTTTCATGACAGATGTTAAACAGCTGGCTGCTC 7462
 7271 CCGGAGCAAAATCAATGATGTGAAGATTACATCAGGATGTTTACAGCTGGCTGGTG 7330
 7463 ACCTTTGATTTCCAGCTACACAACTGATCTCGTGTATCCAAACCCAGACATGATGCC 7522
 7331 ACGTTTGGCTTCCATCTGCACAAATGCTATTCCTGATTCCTGTTCCTCCAAATTTGATTTA 7390
 7523 ATGGAACCTCTACAGGCTCATCCACACAGATGAAACCCAGAGGTGGGACACAGC 7582
 7391 ACTGAGCTTCTCTATGAGCT-----TGGAAGAGTCAACAGTGGGAAGATGTG 7438
 7583 AAGTCTATCTCTCGGGTACAGTGTGAAGTACAGAAAGCAGCTCAAGGCTTTGTCACTTTA 7642
 7439 CCGCCCATCTTTGGAGTTCAGCAGCAAGTGGCAAGGCAAGCCATCTTGTGCTCTG 7498
 7643 GAACGTTTGAACAGCTTATGGCTCCAAATCAACAGCTGCCAGCAGCTCCAAAGAC 7702
 7499 GGAAGATGGCGAGGTGACAGTGGCGGACGACCAAGCTGGCGCGAG---CAGTCTGTGG 7555
 7703 AAGAAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTTGAAGAT 7762
 7556 CTGTGTTTCGACACGGTCAAGTCTCATCGCAAGGGCTCATGCTGGCGCTGAGCCAA 7615
 7763 GCGGAGTGACCAACAGATCATAGTGTGGCCAATGAGGATGGCGGAAGGTTGCTGCC 7822
 7616 GCGCGCTGCAGACCAACGCTCAACATCGCAACGAGGACTGCATCAAGGTGGCGGG 7675
 7823 ATCTTGAACCATGCCCACCTACCTAGAGACCTGCACTTCCACATGATGGGTGATACC 7882
 7676 GTGCTCAACACGCTTCTTACCTGGAGAACCTGCACTTCCACATCGAGGCAAGACACA 7735
 7883 CATTAATTTGTAAACCCAGGACCTTCAGAAAGTGCACCTGGCCATCTCGGCTCAGTGGG 7942
 7736 CACTACTTCATCAAGACCAACACACCGAGAGCGACCTGGGCACTCGCGCTGACGAGC 7795
 7943 GGGCGGCAACCTCGGAGATGGGTCAACGTCACTGTGCTCCAGATCAACACAGTACTT 8002
 7796 GGTGCAAGGCGCTGGAGAACCGGATCAACGTGACCGTGTCTCAGTCCACACGCGTGGT 7855
 8003 AATGGCAGGACTAGACGCTTACACAGATCCAGCTCCAGTACGGGGCACTGTGCTTTGAAC 8062
 7856 AACCGCAGACTCTCGAGGTTCCCGACCTGGAGATGCACTTGGGTGCTCGTGCATGCA 7915
 8063 ACACGCTACGGGACAACTTTGGATGAGGAAGGACAGGCTCTGAGCTGCCCGGAG 8122
 7916 GTGCGCTATGGGATGACGCTGACGAGGAGAGGCGCGCAATCTGGAAGCAGCGCGCGCAG 7975
 8123 AGAGCGCTGCGCAACGCTGGCGCGGCGAGCAGAGACTGCGGGAAGGGAGGAGAACGC 8182
 7976 CCGCGCTCGCCCGGCGTGGGACAGGAGCAGCGGCTGCGGACGCGGAGAGGGT 8035

QY	1526	GATGCGAGGAGCTCTTAACCCAGGAGCGCGGAGCCTAGAGGGACCCCGCCAGTCT	1585	DB	2578	TGCATGGATCCCGATTGCTGCTACAGAGTTCTGCCAGAATCAGCCCTATTGTGGGGA	2637
DB	1498	GATGCGAGGAGCTGATTGCCAGAGAGCAGCGGAGCTGCTTGAGACGGAGAGCGGG	1557	QY	2648	TCCCTTAACCTCTCGACATCATCCAGGAGACAGAGTCCCTGTGTACACAGAACCTA	2707
QY	1586	CGGGAACTGTGCCCCCTCCAGCCATGAGACAGGCTTCATCAGTATTTGGATTGAGGA	1645	DB	2638	CTGCCGATCTCTCAGGACATCATTAGCCAAAGCCTTCAATCGCTTCTCAGCAAGCTGCC	2697
DB	1558	CGGAGGCGAGATCCGTGAGCTTCATGAGCCCGCTTTATCCAGTACTTGGATTCTGGA	1617	QY	2708	CATCTCTTTATGACCGCATCAAGTTCTCTGTCGSCAGGACAGCAGCACAATAATCCCC	2767
QY	1646	ATCTGGCACTTGGCTTTTACAAATCAGGAAAGAGTCAAGAGTGGTTCTCTTCTCACCC	1705	DB	2698	AAATCCTTTATGATCGAATCAGTTCTCTATAGCATCTAGATAGCAGCCATGTATACCT	2757
DB	1618	ATCTGGCATCTGGCTTTTATTAATGATGGGAAATGCAGAGCAGGTGTCTTTTATACC	1677	QY	2768	GGGAGAGACCCCTTTGATGGAGGAGCATCTTGTGTATTCTGTCGCAAGTATGATCATCA	2827
QY	1706	ACTGCCATTGATGCTGGTGAATCACTGCCCCAGCAACTGCTATGGCAATGGTACTGCATC	1765	DB	2758	GGAGAAAGTCTCTTCAATAAGAGCCTTGCACTCTGTCATCAGAGGCCAAGTACTGACTGCT	2817
DB	1678	ATTGTATTAGAGTCTGTGGTGAATGTCCCGCAATTTGCCATGGAAATGGAGAAATGCGTT	1737	QY	2828	GATGGAACCCCTCTGTTGGTGAACATCAGTTTGTCAATAACCCCTCTCTTTGGATAT	2887
QY	1766	TCTGGCACTGCCACTGCTTCTGGTTCCTGGGCTTCTGGGCCCCGACTGTGCGAGAGCTCTGCG	1825	DB	2818	GATGGAATCTCCACTTATTTGGAGTAAATGCTCTCGTTTTTCCATTACCCAGAAATATGATAT	2877
DB	1738	TCTGGAATCTGCCATTGTTTTCAGGATTTCTGGCTCGGATTTGTTCAAGAGCGGCTGT	1797	QY	2888	ACAATCAGAGGACCAAGATGGCAGCTTTGACTTTGGTGACAAATGGCGCATCTCCATCATC	2947
QY	1826	CCGTGCTCTGTAGCGGAAATGGCCAAATACATGAAAGGACAGATGCTTTGTGCCAGTGGC	1885	DB	2878	ACTATTACCCCGCAGGACGGAATGTTGACTTTGGTGCAAAATGGTGGGCGCTCTCTAACT	2937
DB	1798	CCAGTGTATTGTAGTGGCAACGGCAGTACTCCAAAGGGCGCTGCTGTGTTTTCAGCGGC	1857	QY	2948	CTGCGGTTTCGAGCGGCGACCTTTTCAACACAGAGGACACCCCTGTGGCTGCCATGGAT	3007
QY	1886	TGGAAGGCGCTGAGTGGATGCTCCACCAACAGTGTATCGATGTTGGCTCCAGCAAC	1945	DB	2938	TGTTGATTTGAAAGCATCCCATTTCTCACTCAGTATCATATGTTGTTGGAATTCATGNAAT	2997
DB	1858	TGGAAGGCAACGAGTGTGATGTCCGACTTACCAGTGTATTGACCCACAGTGTGGGGGT	1917	QY	3008	CGCTTCTTTGTGATGGAACCATCATATGAGACATGAGGAGATGAGATTTCCAGCTGT	3067
QY	1946	CATGCGACTGCATCAGGCGACCTGCTATCTGCAACCTCTGGTACAAAGGGCGAGAGCTGT	2005	DB	2998	GTCTTTTATGTGATGGATACCTTAGTCATGGAGAAAGAGAAATGACATTTCCAGCTGT	3057
DB	1918	CGTGGATTTGTATATGCGGCTCTGCTGCTTGCAGCTCAGGATACAAAGGAGAAATGTT	1977	QY	3068	GACCTGACCAATTTTCCCGCCGCCAACCCAGCTGCTCTCCATCCCACTGACCTGCTTC	3127
QY	2006	GAGGAAGTGGACTGATGAGACCCACATGTTTCAGGCGGGGTGCTGCTGAGAGGCGAA	2065	DB	3058	GATCTGAGTGGATTCGTGAGGCCAAATCCCATCATTTGTTGTCATCACCTTTATCCACCTTT	3117
DB	1978	GAAAGAGTGAATGTATAGACCTTGGGTGTTCTAATCATGTTGTTGTTATCCACGGGAA	2037	QY	3128	GCCAGCTCTCTGTCAGAGAAAGGGCCCATTTGTCGGGAAATTCAGGCTTTTCAGAGAGAA	3187
QY	2066	TGCCATTTCTTGTGGATGGGAGGACCAACATGCGAGACCCCGAGGCGCATGCTTA	2125	DB	3118	TTTCAGATCTCTCTCTGAAAGACAGTCCCATCATTTCCGAAACACAGGTACTCCACAGGAA	3177
DB	2038	TGTCAGTCACTCCAGATGGGAGGTAGCAATTTGTGAATATCTGAAGACATGTGTCCA	2097	QY	3188	ATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCGGACCCCTGGCTAC	3247
QY	2126	GACCATGTTTCAAGGCAAGAACCTTCTCCCGGACACCGGGCTTTGACGTGTGACCCA	2185	DB	3178	ACTACAATTTCCAGGAACAGATTTGAAACTCTCTCTACTTGGTTCCAGAGCTGCAAGGTAT	3237
DB	2098	GACCATGCTCTCCGGCCAGCAACATATCTTCAAGAAAGTGGCTCTGACGTGTGACCT	2157	QY	3248	AAATCTGCTCTGAGGATCAGCTCACCCACCGGACCATCCCTTCAACCTCATGAAGGTG	3307
QY	2186	AGCTGGACTGACACGACTGTTCTATCGAGATCTGCTGCCAGCTGTGTGGCTGCTGCG	2245	DB	3238	AGTCACTTCTCAAGATCACCATGACCCAGTCTATTATTCCATTTAATTAAGGTT	3297
DB	2158	AACTGGACTGCCAGACTGCTCAACAGAAATATGTTCTGTGACTGTGCTCACAGGC	2217	QY	3308	CACCTCATGTTAGCGGTGGAGCGGCGCTTTCAGGAAGTGTGCTGCTGAGCCCGACAG	3367
QY	2246	GTGTGCTGAGGGGCACTGCGCTGCGAGGATGGCTGGATGGGGGAGCGCTCGACAG	2305	DB	3298	CATCTTATGGTGTAGTAGGAAGACTCTTCCAAAGTGGTTTCTCTGCTCCACCAAC	3357
DB	2218	GTGTCATGGGGGAGCGTGTGCTGTGAAGAGGCTGGACGGGCGCCAGCTGTAATCAG	2277	QY	3368	CTGCTCTTATTTTCATTTGGGCAAGACAGACGCTCTACCAACAGAGGTCTTTGGGCTT	3427
QY	2306	CGGCGCTGCCACCGCTGTGCGAGCATGGGACCTGCGCGGACGCGAGCTGCGAGTGC	2365	DB	3358	TTGGCTATATCTTTCATATGGATTAACAGATGCATATATAGAAAGTCTATGGTCTA	3417
DB	2278	AGGCGCTGCCACCGCTGTGCGGAGCAGCGGACCTGCAAGATGGCAAGTGTGAATGC	2337	QY	3428	TCAGAAAGCTTTGTTTCCGTGGGTTATGAATATGAATCTCTGCCAGATCTAATCTGTGG	3487
QY	2366	AGCCCTGCTGGAATGGCAACACTGCAACCATGCTCACTATCTGGAAGGAGTGTGAGT	2421	DB	3418	TCGAAAGCTGTTGTGCTGAGTTGGATATGAGTATGAGTCTGTTTGGACCTGACTCTGTGG	3477
DB	2338	AGCCAGGCTGGAATGGAGGACACTGCACTATCGCTCACTATTGGATAGATAGTAA	2397	QY	3488	GAANAAGAACAAACAGTCTCAGCGGCTATGAATTCAGCGCTCCAGCTTGGAGGATGG	3547
QY	2422	-----TAAAGGGTGGCTGGTGTGCAATGGCAACGCGCATGTACC	2467	DB	3478	GAANAAGAGGACTGCAATTTCTGAGGCTTATGAATGGATGGTCCACATGGGTGGCTGG	3537
DB	2398	GACAAGATAGGATATAAGAGGTTGCTGTGCTGTGCAACAGCAATGGAAGATGTACC	2457	QY	3548	AGCTTAGACAAACATCATGCTCCCTCAACATTTCAAAGTGTATCTCTGCACAAAGGAATGG	3607
QY	2468	TTAGACTGGAATGTGCTGCACTGCTGCGAGCTGGCTGGAGAGGAGTGTGCTGTGAC	2527	DB	3538	ACATTTAGATAAATCATCAGTCTGGAATGATACAGAACGATATCTGTACAGGGAACCGG	3597
DB	2458	CTGACCAAAATGCGGACATTTGTGTGCTGCGCTGGATGGAGGAGCAGCTGTGAC	2517	QY	3608	GAGAACAGATTTGTGCTCAGCAGCCTCTGTATTTGGGACATCATGGGCAATGGGCGC	3667
QY	2528	ACTTCCATGGAGATGCTGCGGTGTACAGCAAGACAAATGATGGAGATGGCTGTGGAC	2587	DB	3598	GAANAAGAGTTCATCTCCAGCAGCTCCAGTCTGTGATGATCATTTGGGCAATGGGCGA	3657
DB	2518	GTAGCCATGGAGACTCTTTTGACAGATAGCAAGGACAAATGAAGGGATGGACTCATTTGAC	2577	QY	3668	CGGAGAACCATCTCTGCTGCCCGCAGCTGCAACCGGCTGCTGACGGCAACAGCTCTCGCC	3727
QY	2588	TGCATGAGCCTGACTGTGCTCCAGCCCTGCTGTGCCATATCAACCGGCTGTGCTTGGC	2647				

Db 3658 AGGCGAGCATTTCTGCCCCAGTTGCAATGGTCAAGTGGTAAACAAGTTACTGGCC 3717
 Qy 3728 CCAGTGGCCCTCAGCTGTGGCTCTGAGGGAGCCTCTATGTGGGTGATTTCAACTACATT 3787
 Db 3718 CCAGTGGCGTAGCTTGTGGGATCGATGGCGAGTGTACGTAGCGGATTTCAACTACGTG 3777
 Qy 3788 AGAAGGATCTTCCCTCTGGAAATGTGACCAACATCTAGAGCTGAGGAATPAAGATTT 3847
 Db 3778 CGGGGATATTCCCTCTCGAAATGTAAACAAGTGTCTTAGAACTAAAGAAATAAGATTTT 3837
 Qy 3848 AGACATAGTCAGTCCAGCAGACAAATACCTACCTGGCCACAGACCCCATGAGTGGGGCC 3907
 Db 3838 AGACATAGCAGCAACCCAGCTCATAGATCTACCTTGCACCGGATCCAGTACCGGAGAT 3897
 Qy 3908 GTCTTCTCTCTGACAGCAACAGCGCGGCTCTTTTAAATCAAGTCCACTGTGGTGGT 3967
 Db 3898 CTGTACGTTTCTGACACAAACACCGCAGAAATTTATGCCCAAAGTCACTTACGGGGCA 3957
 Qy 3968 AAGGACCTTGTCAAGAACTCTGAGGTGTGGGGGACAGTGCAGTGCCTCCCTTT 4027
 Db 3958 AAGACTTGTATTAATAATGCAAGATGTCTGACGGGACAGGGGAGCAATGCTTCCGTTT 4017
 Qy 4028 GATGACATCTCGCTGGGGATGTGGGAAGCCACAGAGCCACACTCAACATCCCGAG 4087
 Db 4018 GACGAGGCGAGATGTGGGGATGGAGGAAGCCGCTGGAAGCCACACTCATGAGTCCCAA 4077
 Qy 4088 GGTATTACAGTGGACAAGTTTGGGCTGATCTACTTGTGGATGSCACCATGATCAGAGC 4147
 Db 4078 GGAATGGCAGTGTGAAGAAATGAATTAATCTACTTGTGATGAACCATGATTAGGAA 4137
 Qy 4148 ATCGATCAGATGGGATCATCTCAACCTGTCTGGCTCTAATGATCTCACTACAGCCCG 4207
 Db 4138 GTTGACCAATGGAATCATATCACTCTTCTGGCTCTACGATTTGACTTTCAGCCAGA 4197
 Qy 4208 CCACTCAGCTGTGATCTGTGATGATATTTCCAGGTAAGACTGGAGTGGCCCAAGAC 4267
 Db 4198 CTTTAACTGTGACACAGCATGCACATCAGCCAGGTACGCTCTGGAATGGCCCACTGAC 4257
 Qy 4268 TTAGCCATCAACCAATGGAACATCTTTATGCTCTGACACAAATGTGCTCTGCA 4327
 Db 4258 CTAGCCATTAACCTTATGGATACTCCATTTATGCTCTGGATAAATGATGATTTACAG 4317
 Qy 4328 ATCTCTGAAAACCAAGTGGCATTTGTCCGGGAGCCCATGCACTGCCAGTCCCT 4387
 Db 4318 ATCACTGAAATCGTCAAGTTGCGATTCGCTGGAGGCCCATGCACTGTGAGTTCCC 4377
 Qy 4388 GGCATGACCACTTCTCTGTAAGCAGAGTGGCCATCCAGCAACCTGGAGTCAAGCACC 4447
 Db 4378 GAGTGG---AATATCTCTGGGAAAGCAGCGGTGCAGACACACTGGAATCAGCCACT 4434
 Qy 4448 GCTTTGGCTGTTTACAAATGGGCTCTGTATATTGCTGAGACTGATGAGAAAAGATC 4507
 Db 4435 GCATTTGTGTCTCAAGTGGGCTCTGTACATTAATGAACTGATGAGAAAGAAAT 4494
 Qy 4508 AACCGCATCAGCAGGTCAACATAGTGGAGATCTCACTGCTGTGGGGCCCCCAGT 4567
 Db 4495 AACCGGATAAGCAGGTCAACACAGATGGAGAAATCTCTTAGTGGCCGGAATACCTTCA 4554
 Qy 4568 GCTGTGACTGTAAATGATGCAACTGATGATGTTTCTGGAGAGATGGTTATGCC 4627
 Db 4555 GAGTGTGACTGCAAAAATGATGCCAACTGTGACTGTTACAGAGTGGAGATGGGTACGCC 4614
 Qy 4628 AAGGATGCAAAAGTAAATACCCCATCTTCTTGGCTGTGTGTGCTGATGGGAGCTTAC 4687
 Db 4615 AAGGATGCCAAACTCAGTGGCCCATCTCTCTGGCTGCTTCTCCAGATGGTACACTGTAT 4674
 Qy 4688 GTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAGCCCTTCTCTAC 4747
 Db 4675 ATTGCAGATCTAGGGAATATCCGGATCCGGCTGTGTCAAGAAATAAGCCCTTTACTTAAC 4734
 Qy 4748 ACCCAAGACATGATGAGCTGTCTTCAACAATGACAGGAGCTCTATCTGTGTTGATACC 4807
 Db 4735 TCTATGAATCTTATGAAGTTCGCTCTCCAATGATCAGAACTCATATCTTTGACATC 4794

Qy 4808 ACCGCAAGCACCTGTGTACACCAAGCCTGCCACAGAGACTACTCTGTACAACTTACC 4867
 Db 4795 AATGTTACTCACCAATATATCTGAAAGTTTAGTCACTGGTGAATACCTTTTACAAATTTAGC 4854
 Qy 4868 TACATGGGAGCGGAGACATCACACTCATCAGACAAACAATGGCAACATGTAATGTC 4927
 Db 4855 TACAGCAATGCAATGATATATCTGCTGTGACAGACGAATGGCAACACCTTTAGAAAT 4914
 Qy 4928 CGCCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGTGTACTGG 4987
 Db 4915 AGAGCGGACCCAAATCGCATGCCAGTTCGAGTGGTGTCTCTGATAAACCAAGTGATATGG 4974
 Qy 4988 GTGACCATGGGACCAACAGTGCCTCAAGAGTGTGACACACAGGACACAGTTGGCC 5047
 Db 4975 TTGACAAATAGGAACAAATGATGTTTGAAGGCATGACTGCTCAAGGATGGAATTAGTT 5034
 Qy 5048 ATGATCACATACCATGGCAATTCGGCCCTTCTGGCAACCAAAAGCAATGAAAACCGATGG 5107
 Db 5035 TTGTTTACTTACATGGCAATAGTGGCTTTTAGCCACTAAAGTGATGAAACTGGATGG 5094
 Qy 5108 ACACAATTTTATGATGACACAGCTTTGGCCGCTGTGACAAATGTGACCTTCTCTACTGGC 5167
 Db 5095 ACAACGTTTTTTGACTATGACAGTGAAGGTCTGTGACAAATGTTACGTTTCCAATGGA 5154
 Qy 5168 CAGGTGAGCAGTTTCCGAAAGTGTATCAGACAGTTCAGTGCATGTCAGGTAGAGACTCC 5227
 Db 5155 GTGGTCACAAAACCTGATGGGACATGGAAGGCTATCAGAGTGAATGAGTCACT 5214
 Qy 5228 AGC---AAGGATGATGTACCAATAACCAACCACTGTCTGCTCAGGCGCTTCTACACA 5284
 Db 5215 AGCCGAGAAAGATGTGACATCACTTCAATCTGTCTCGATCGATTTCTTCTACACC 5274
 Qy 5285 CTGCTCAAGACCAAGTCCGGAACAGTACTACATCGGGCCGATGGCTCTTGGCGCTG 5344
 Db 5275 ATGGTTCAAGATCAGTTAAGAAACAGCTACCAAGTTGTTATGACGGCTCCCTCAGAAAT 5334
 Qy 5345 CTGCTGGCCAAACGGCATGGAGTGGCGCTGCAAGCTGAGCCCACTTGTGCTGGTGGCACC 5404
 Db 5335 ATCTAGCCAGTGGCTGCACTCACACTACCAACAGAGCCGCACTTCTGGCTGGCACC 5394
 Qy 5405 GTCAACCCCAACCGTGGGCAAGGAATGTGACGCTGCCCATGACAAAGGCTCAACCTG 5464
 Db 5395 GCTAATCCGACGTTGCCAAAAGAAACATGACTTTGCTGGGAGAACGCTCAAAACTTG 5454
 Qy 5465 GTGAGTGGCGCCAGCGCAAGAGAGGCTCGGGGCGAGTCACTGCTTTGGGCGCCCG 5524
 Db 5455 GTGGAATGGAGATTCGAAAAGAGCAAGCCCAAGGAAAGTCAATGTCTTTGGCCGCAAG 5514
 Qy 5525 CTGCGGTGACAAACCGAAATCTCTTATCTGTGACTTTGATGCGGTAAACGCAAGAG 5584
 Db 5515 CTCAGGGTTAATGGCAGAAACCTCTCTTCAAGTTGACTTTGATCGAAACAAAGACAGAA 5574
 Qy 5585 AAGATCTATGATGACACCGCAGTTTCACTTCCGATTTCTGTACGACCGAGGCGCG 5644
 Db 5575 AAGATCTATGACACCGTAAATTTCTTACGAGGATCGCTTACGACAGTCTGGGCAC 5634
 Qy 5645 CCCAGCCTCTGTGTCACCCAGCAGCGCTGAATGGTGTCAAGTGTACATACTCCCCCTGGG 5704
 Db 5635 CCAGCTCTGTGCTGCCAGCAGCAAGCTGATGGCCCTCAATGTCACTTATTCATCCACA 5694
 Qy 5705 GGTTAATTGCTGGCATCCAGAGGGGCATGCTCTGAAAGAAATGGAATACGACAGGCG 5764
 Db 5695 GGTCAAAATGCCAGCATCCAGCGAGGACCACTAGCGAGAAAGTAGATTATGACGGAACAG 5754
 Qy 5765 GCGCGCATCATCCAGGATCTTCTGCTGATGGGAACATGGAAGTGTACATATCTTAGAG 5824
 Db 5755 GGGAGATCTGTCTCGGCTTCTTGTGATGTTAAACATGAGTTTACATATTTAGAA 5814
 Qy 5825 AAGTCCATGCTGTCTACTACACCGCAGGAGCAGTATCTTTGAGTTGCAAGAAAT 5884
 Db 5815 AAGTCCATGCTTCTTCTGCTTCTATAGCCAGCGCAGTACATCTTTCGAATACGATATGTTGG 5874

Db	8017	GGCGCGCTGGCGCTGCACGTGCGCTACGGCATGACCCCTGGACGAGGAAGCGCGCATC	8076
Qy	8105	CTGGAGCTGGCGCGGAGAGACCGCTGGCCAAAGCGTGGCGCCCGGAGCAGCAGACTG	8164
Db	8077	CTGGACAGCGCGGAGCGCGCGCTGCGCCGGGCTGGCGCGGAGCAGCAGCGGTG	8136
Qy	8165	CGGGAAGGGAGGAAGCGCTGGCGGCTGGACAGAGGGGGAGAGCAGCAGGTCTGAGC	8224
Db	8137	CGCGACGGCGAGAGGGCGCGCGCTCTGGACGGAGGGCGGAAGCGCGCAGCTCTGAGC	8196
Qy	8225	ACAGGGCGGGTGCAGAGCTACCAAGCGCTTTTTCGATCTCTGTGAGCAGTACCCAGAA	8284
Db	8197	GCGGCAAGGTGCAGGGCTACAGCGGTACTAGTACTCTCGTGGAGCAGTACCCGAG	8256
Qy	8285	CTGTACAGAGCGCCCAACATCACTTTCACTAGACAGAGCGAGATGGCGCGAGTGA	8344
Db	8257	CTGGCGAGAGCGCCAAACATCCAGTTCTCTGGCGAGAGCGAGATCGGAGAGGTAA	8316
Qy	8345	C 8345	
Db	8317	C 8317	
RESULT 11			
AX662353			
LOCUS	AX662353	8675 bp	DNA linear PAT 22-MAR-2003
DEFINITION	Sequence 35 from Patent WO02062999.		
ACCESSION	AX662353		
VERSION	AX662353.1	GI:29163216	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1. Spytek, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X., Malyan, K., Shinkels, R.A., Tchernev, V.T., Spaderna, S.K., Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangoli, E.A., Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F., Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G., Millet, I., and Macdougall, J.R.		
TITLE	Proteins and nucleic acids encoding same		
JOURNAL	Patent: WO 02062999-A 35 15-AUG-2002; Curagen Corporation (US)		
FEATURES	Location/Qualifiers		
source	1..8675		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	41.6%;	Score 3473.8;	DB 6; Length 8675;
Best Local Similarity	65.2%;	Pred. No. 0;	
Matches 5445;	Conservative 0;	Mismatches 2692;	Indels 216; Gaps 14;
Qy	35	ATGAGCTGAAGAGAGGAAGCCTTACCGCTCGCTGACCCGCGC---CGCGACCGGAG	91
Db	151	ATGGAATGAAGAAAGCGAGCGCTTACTGCTCCCTGACCAAGAGCAGCAGAGAAGGAA	210
Qy	92	CGCGCTACACCAAGCTGTCGCGGACAGCAGGAGGGCAAG---CCCGCAGAAATCG	148
Db	211	CGGCGCTACACAAATTCCTCCGAGACAAATGAGAGTGCAGGATACCCACAGAGTCC	270
Qy	149	TACAGCTCCAGCAGACCCCTGAAGGCGCTACGACCCAGGA---CGCCGCGCTAGCCTATGCC	205
Db	271	TACAGTTCCAGCAGACATTTAAAGCTTTTGAATCATGATTCCTCGCGGCTGCTTTACGGC	330
Qy	206	AGCGCGTCAAGACATTTGTCGCGAGGAGCGCGAGGAATTCGCGCAGAGCTGCCAAC	265
Db	331	AACAGAGTGAAGATTTGGTTTCAGAGAGCAGACAGAGTTCCTACTAGACAGGACAGAA	390
Qy	266	TTACCCCTCGCGGAGCTGGGCGCTGGAGGAAGTAACCCCGCTCAAGGGACCCCTGTACCGG	325

QY 1406 ATAGACCATCTGTGTCATCTGAAATTCATATGTTCTCTGGGAAAGCGCAGCCCTGGTTGGC 1465
 Db 1378 ATTGATCAGCCACAGTTTCTTAAATTCATATCTCTCTCAAGAGATGCAATTGATTGA 1437
 QY 1466 ATTTATGGCAGAAAAGGCTCCCTCTTACATA-----CACAGTTGACTTTT 1513
 Db 1438 GTATATGGCGAAGATTACCGCTTCCCACTCAGTCTCCCGCCAGTATGACTTC 1497
 QY 1514 GTGAGCTGTGATGGCAGAGGCTCCTAAACCCAGGAGGCGGAGCCTAGAGGGAC 1573
 Db 1498 GTGAGCTCTCTGATGGCAGAGGCTGATTCAGAGAGCAGCGAGCCTCTTGAAG 1557
 QY 1574 CCGCCCGAGCTCTGGGAACTGTGCCCTCCAGCCATGAGCAGGCTTCATCCAGTAT 1633
 Db 1558 GAGAGCGCGGCGCAGCGAGATCCGTACGCTTCATGAGCGCGCTTTATCCAGTAC 1617
 QY 1634 TTGGATTGAGGAATCTGGCATTGGCTTTTACAAATGACGAAAGAGTCAAGAGTGGTT 1693
 Db 1618 TTGGATTGGAATCTGGCATTGGCTTTTATAATGATGGAATAATGACAGCAGGTG 1677
 QY 1694 TCCTTTCTCACCAGTCCATTGAGTCGTGATTAATCCCGCAGCACTGCTATGGCAAT 1753
 Db 1678 TCTTTTAATACCAATGTTATAGATCTGTGTGAAATGTCCCGAAATGTCATGGAAT 1737
 QY 1754 GTGACTGCATCTCTGGACCTGCGACCTGCTTCTGGGTTTCTGGGCCCGCAGCTGTGC 1813
 Db 1738 GGAGAAATGGTTCTGGAACTTGGCAATGTTTTCAGGAATTCGGGTCCGGATTGTCA 1797
 QY 1814 AGAGCCTCTGCGCGCTCTGTAGCGGAATGCGCAATACATGAAAGCAGATGCTTG 1873
 Db 1798 AGAGCGCGCTCTGAGTCTTATAGTGGCAACGGGAGTACTCCAAGGCGCGTGGCTG 1857
 QY 1874 TGCCACAGTGTGGAAGGCGCTGAGTGCATCTGCCACCAACAGTGTATCGATGTG 1933
 Db 1858 TGTTTACGCGCTGGAAGGCGCAGTGTGATGTCGCGACTCCAGTGTATGACCCA 1917
 QY 1934 GCCTGACAGCAATGACCTGTCATCAGCGGCACTGTCATGCAACCTCGGCTACAAG 1993
 Db 1918 CAGTGTGGGGTCTGGGATTGTATCATGGCTCTGTGCTGCTCAGCTCAGGATACAA 1977
 QY 1994 GCGAGAGCTGTGAGGAGTGGATGTCATGACACCCACATGTTTCAGCGCGGTGTGTC 2053
 Db 1978 GGAGAAAGTGTGAAGAGTGACTGTATAGACCTCGGTGTCTAATCATGTGTGTGT 2037
 QY 2054 GTGAGAGCGGAATGCCATTGCTTTGTGGATGGGAGGCGCAACTCGGAGACCCCGAG 2113
 Db 2038 ATCCACGGGGAATGTCACTGAGTCCAGGATGGGAGGTAGCAATTTGAAATACTGAAG 2097
 QY 2114 GCCACATGCTTAGACAGTGTTCAGGCAAGCACTTCCTCCGCGACACCGGCTTTGC 2173
 Db 2098 ACCATGTGTCCAGACCACTGCTCCGCGCACGGAACTGCTTCAACCGAAATATGTTCTGTGAGTGT 2157
 QY 2174 AGCTGTGACCAAGCTGAGTGGACACAGCTGTTCTATCGAGATCTGTGCTGCGACTGT 2233
 Db 2158 ACGTGTGACCTTAAGTGGACTGGCCAGACTGCTCAACCGAAATATGTTCTGTGAGTGT 2217
 QY 2234 GGTGGCAGTGTGCTGAGGAGGCACTGCGCTGCGAGAGTGGCTGGATGGGAGGCA 2293
 Db 2218 GGCTCACACGGGTTTGTATGGGGGGGACGTGTGCTGTGAAGAAAGTGGCTCCTGTC 2277
 QY 2294 GCTGTGACACGAGGCGCTGCACCGGCTGTGCGGAGCATGGGACCTGCGCGAGCGC 2353
 Db 2278 GCTGTATCAGAGAGCTGCCCCCGCTGTGCGGAGCAGGAGCTTCAAGAGATGGC 2337
 QY 2354 AAGTGGAGTGCAGCCCTGGCTGAATGGGAACTGCACTGCACTGCTCACTATCTGGAT 2413
 Db 2338 AAGTGTGAATGCAGCCAGGCTGGAATGGAGAGCACTGCACTATCGCTCACTATTTGGAT 2397
 QY 2414 AGGTGACT-----TAAAGAGGTTGCCCTGGTGTGCAATGGCAAC 2455
 Db 2398 AAGATGTTAAAGACAGATAGGATATAAAGAGGGTTGCTCTGCTGTGCAACAGCAAT 2457

QY 2456 GGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGA 2515
 Db 2458 GGAAGATGTACCTTGGACCAAAATGGCGACATTTGTGTGCGCAGCTGGATGGAGAGA 2517
 QY 2516 GCTGGCTGTGACACTTCCATGAGAGCTGCTCGGTGACAGCAAGACAATGATGGAGAT 2575
 Db 2518 GMAGGCTGTGAGTGTAGCATGGAGACTCTTTGACAGATAGCAAGACAATGAAGGGAT 2577
 QY 2576 GGCCTGGTGGACTGTGACCTGACTGCTGCTCCAGCCCTGTGCCATATCAACCCG 2635
 Db 2578 GGAATCATTTGATGATGATCCGATCCGATTTGCTGCTACAGAGTTCTGCCAGATCAGCC 2637
 QY 2636 CTGTGCTTGGCTCCCTTAACCTCTGAGCATCATCCAGGAGACACAGTCCCTGTGTCA 2695
 Db 2638 TATGTGCGGGACTGCGGGATCTCAGGACATCATTAGCAAGCTTCAATCGCCTTCT 2697
 QY 2696 CAGCAAACTTACACTCTCTTATGACCGCATCAAGTTCTCTGGGCGAGGACAGCAGC 2755
 Db 2698 CAGCAAGCTGCCAAATCTTTTATGATCGAATCAGTTTCTTATAGGATCTGATAGCACC 2757
 QY 2756 CACATTAATCCCGGGAGAAACCCCTTTGATGAGGAGCATGCTTGTGTTATTTGCGGCCAA 2815
 Db 2758 CATGTTATACCTGGAGAAAGTCTTTCAATAAGAGCCTTGCATCTGTCTATCAGAGCCAA 2817
 QY 2816 GTGATGACATCAGATGGAAACCCCTGTTGGTGTGAACATCAGTTTGTCAATAACCTCT 2875
 Db 2818 GTACTGACTGTGATGGAACTCCACTTATGAGGTAAATGTCTGTTTTCAATTACCCA 2877
 QY 2876 CTCCTTGGATATPACAAATCAGCAGGCAAGATGGAGCTTTGACTTTGGTGACAAATGGCGC 2935
 Db 2878 GAATATGATATACATATTACCGCGCAGGACGAAATGTTGACTTGGTGGCAATGGTGG 2937
 QY 2936 ATCTCATCATCTCGGCTTGGAGGGGACCTTTCATCACACAGGAGCACACCTGTGG 2995
 Db 2938 GCCTCTTAACTTTGGTATTTGAACGATCCCATTTCTCATCTAGTATCATACTGTGTGG 2997
 QY 2996 CTGCCATGGATCGCTTCTTTGTCTATGGAAACCATCATCATGAGACATGAGAGAAATGAG 3055
 Db 2998 ATTCCATGGAATGTCTTTTATGTGATGATACCTCTAGTATGAGAAAGAGAGAAATGAC 3057
 QY 3056 ATTCACAGCTGTGACTGAGCAATTTTCCCGCGCCCAACCCAGCTGCTCTCCATCCCA 3115
 Db 3058 ATTCACAGCTGTGATCTGAGTGGATTCGTGAGGCGAAATCCCATCATTTGTCTCATCCT 3117
 QY 3116 CTGAGCTCTTCCGCGAGCTCTCTGTCAGAGAAAGGCGCCCATTTGTCGCGAAATTCAGCT 3175
 Db 3118 TTATCCACTTTTTCAGATCTTCTCCTGAAGACAGTCCCATCATTCGCGAAACAGGTA 3177
 QY 3176 TTGCGAGGAGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCGG 3235
 Db 3178 CTCCACGAGGAATCTCAAAATTCAGGAACAGATTTGAACCTCTCTACTTGTGATTCAGA 3237
 QY 3236 ACCCTGTGCTCAAAATCTGCTGAGGATCAGCTCACCACCGCAGCATCCCTTCAAC 3295
 Db 3238 GCTGAGGAGTATGATGATGTTCTCAAGATCACCAGTCCAGTCTATTTATTTCAATTAAT 3297
 QY 3296 CTCATGAAGTGCACCTCATGATGCGGTGAGGCGCGCTCTTTCAAGAAATGGTTCCGT 3355
 Db 3298 TTAATGAAGTTTCTCTTATGTTAGTGTAGTGTAGGAGACTCTTCCAAAGTGGTTTCT 3357
 QY 3356 CGAGCCCGCAGACCTGTCTTATTTTCAATTTGGGACAGAGCAGCTCTACAAACAGAG 3415
 Db 3358 GCCTCACCAAATTTGGCTCTACTTTTCAATTTGGGATAAACAGATGATATATACAGAA 3417
 QY 3416 GTGTTTGGCTTTTCAAGAGCCTTTTCTTCTCGTGGGTTTATGAATATGAATCTGCCCCAGAT 3475
 Db 3418 GTCTATGTTCTATCTGAAGCTGTTGTGTCAGTTGATGATGATGATGATGTTGTCAG 3477
 QY 3476 CTAATCTGTGGGAAAAAGAACCAACAGTGTGAGGCTATGAATTTGAGCGTCCAG 3535
 Db 3478 CTGACTCTGTGGGAAAAAGAGACTGCCATTTCTGAGGGCTATGAATTTGGATGGTCCAGC 3537
 QY 3536 CTTGGAGGATGGAGCTAGACAAACATCATGCGCTCAACATTCAAAGTGGTATCTCTGCAC 3595

Db	3538	ATGGTGGCTGGACATTAGATAAATCATCGTGTCTGATGTACAGACGGTATATCTGAT	4615	GATGGCTAGCCCAAGGATGCCAACTCAGTGGCCCATCTCTCTGGCTCTCTCCAGAT
Qy	3596	AAAGGGAATGGGGAAGAACAGTTTGTCTCAGAGAGCTCTCTGATTCATGGAGCAATCATG	4676	GGGGAGCTCTACGTGGCCGACCTTCGGGAACATCCGAATTCGGTTTATCCGGAAGAACAG
Db	3598	AAGGGAACCGGGGAACCAATTCATCTCCAGAGAGCTCCAGTCGTGAGTACATCATG	4675	GGTACACATGTATTGTGAGATCTAGGGAATATCCGGATCCGGCTGTGTCAAGAAATAG
Qy	3656	GGCAATGGGCGCGGAGAGATCTCTGCCCCAGCTGCAACGGCCCTTGTGACGGCAAC	4736	CTTTCTCTCAACCCAGCAACATGTATGAGCTGTCTTCAACAAATGACAGAGAGTCTAT
Db	3658	GGCAATGGGCGAAGGGCGAGATTTCTGCCCCAGTTGCAATGGTCAAGCTGATGGTAAC	4735	CTTTTACTTAACTCTATGAATCTTCTGAAAGTTGGCTCTCCAACTGATCAAGAACTCTAC
Qy	3716	AAGTCTCTGGCCCAAGTGGCCCTCACCTGTGGCTCTCAGCGAGAGCTCTATGTGGGTGAT	4796	CTGTTTGATACCAACGGCAAGCACTGTACACCCAAAGCCCTCCACAGGAGACTACCTG
Db	3718	AAGTTACTGGCCCAAGTGGCCCTAGCTTGTGGATCGATGCGAGCTGTACGTAGGCGAT	4795	ATCTTTGACATCAATGGTACTCACCAATATCTGTAAGTTTATGCTGATGATACCTT
Qy	3776	TTCAACTACATTAGAAGATCTTCCCTCTGGAAATGTCAACCAATCTTAGAGCTGAGG	4856	TACAACTTCACCTACACTGGGAGCGGACATCACACTCATCACAGACAAACAAATGCAAC
Db	3778	TTCAACTACAGTGGCGGAGATTTCCCTCTGGAAATGTAAAGATGTCTTAGAACTAAGA	4855	TACAAATTTTAGCTACAGCAATGACAATGATATTACTGTGTGACAGACAGCAATGGCAAC
Qy	3836	AATAAAGATTTACAGATAGTCAAGTCCAGCACACAAATACTACCTGGCCACAGACCCC	4916	ATGCTAAATGTCCGGGAGACTCTACTGGGATGCCCTCTGCTGGTGGTCCCAATGGC
Db	3838	AATAAAGATTTTAGACATAGCAGCAACCCAGCTCATAGATACTACCTTGCAACGGATCCA	4915	ACCCTTAGAATTAGACGGGACCAATCGCATGCGGAGTTCGAGTGGTGTCTCTGATAAC
Qy	3896	ATGAGTGGGCGGCTTCTTCTTCTGACAGCAACAGCCGGGCTCTTAAATCAAGTCC	4976	CAGGTGTACTGGGTGACCATGGCCACCAACAGTGCACCTCAAGAGTGTGACCAACAAGGA
Db	3898	GTACGGGAGATCTGTACGTTTCTGACACAAACACCCGCAATTTATCGCCCCAAGTCA	4975	CAAGTATATGTTGACATAGAACAAATGATGTTTGAAGGCAATGACTGCTCAAGGA
Qy	3956	ACTGTGTGGTGAAGGACCTTGTCAAGAACCTCTAGGTGGTGTGGGGACAGGTGACCG	5036	CACGAGTTGCCCATGATGATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAAT
Db	3958	CTTACGGGGGCAAAAGACTTGACTTAAATTCAGAGTCTGTCGGCAGGACAGGGGAGCAA	5035	CTGGAATTAGTTTGTGTTTACTTACCATGGCAATAGTGGCTTTTAGCCACTAAAAAGTAT
Qy	4016	TGCTCTCCCTTTGATGACATCGCTGGGGATGGTGGGAAGGCCACAGAGCCACATCTC	5096	GAAGCGGATGACACATTTTATGATGACGACAGCTTTGGCGGCTGACAAATGTGACC
Db	4018	TGCTCTCCCTTTGACGAGCGAGATGTGGGATGGAGGAGGCGCTGGAGCCACATCTC	5095	GAACCTGGATGACAAACGTTTGTGACTATGACGTGAAGGTGCTGACAAATGTTAGC
Qy	4076	ACCAATCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGCACC	5156	TTCCCTTACTGGCCAGGTGAGCAGTTTCCGAAGTGTATACAGACAGTTCAGTGCATGCCAG
Db	4078	ATGAGTCCCAAGGAATGGCAGTTGATGAAGAAATGAAATTAATCTACTTTGTTGATGAACC	5155	TTTCCAACCTGGAGTGTGACAAACCTGCATGGGACATGGACAGCAAGCTATACAGTGGAC
Qy	4136	ATGATCAGACGATCGATFAGAATGGGATCATCTCCACCTGCTCCGCTCTAATGATCTC	5216	GTAGAGACCTCCAGC---AAGATGATGTACCAATACCAACCAACCTGTCTGCCCTCAGGC
Db	4138	ATGATTAGGAAAGTTGACCAAAATGGAATCATATCACTCTCTCGGCTCTTAACGATTTG	5215	ATTGAGTCTATAGCCGAGGAAGATGTGACATCACTTCAAAATGTCTCTCGATCGAT
Qy	4196	ACATCAGCCCGCCACTCAGCTGTGATTTCTGTCTATGATATTTCCAGGTAAGACTGGAG	5273	GCCTTCTACACACTGTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGCGGATGGC
Db	4198	ACTTCAGCCAGACCTTTAACTTTGTGACACCAAGTGCACATCAGCAGGTACGTCGGAA	5275	TCCTTTCTACCATGTTTCAAGATCAGTTAAGAAACAGCTACAGATTCGTTATGACGGC
Qy	4256	TGGCCACAGACTTAGCCATCAACCCAAATGGACAACTCACTTTATGTCTCTCGACAAACAAT	5333	TCCTTGGGCTGCTGTGCTGCCAACCGCATGAGAGTGGCGCTGCAGACTGAGGCCCACTTG
Db	4258	TGGCCCACTGACCTAGCCATTAACCCCTATGATTAACCTCATTTATGTCTCTGATTAAT	5335	TCCTCAGAAATTAATCTACGCCAGTGGCTGGACTCACACTACCAACAGAGCGCGCACTGT
Qy	4316	GTGGTCTGCAAACTCTGAAACCAACAGGTGGCANTTTCGGGGAGGCCCATGGAC	5393	CTGGCTGGCACGCTCAACCCCGCTGGGCAAGAGGAATGTCACTGCTGCCCTCGACAAC
Db	4318	GTAGTTTACAGATCACTGAAATCGTCAAGTTGCGCATTTGCTGTGACCGGCCCATGCAC	5395	CTGGCTGGCACCGCTTAATCCGAGCGTTGCCAAAGAAACATGACTTTTGGCTGGCGGAGAC
Qy	4376	TGCCAGGTCCTGGCATTTGACCACTCTCTGTAGCAAGTGGCCATCCACGCAACCTG	5453	GGCTTCAACCTGTGTGAGTGGGCGCCAGCGCAAGAGCAGGCTGGGGCCAGGTCACTGTG
Db	4378	TGTCAAGTTCGCGAGTGG---AATATCTGTGGGAAGACAGCGGTGCGACAAACACTG	5455	GGTCAAAACTTGTGTGGAATGGAGATTCGGAAGAAAGCAAGCCCAAGGAAAGTCAATGTC
Qy	4436	GAGTCAGCCACCGCTTTGGCTGTTTACACAAATGGGTCCTGTATATTTGCTGAGACTGAT	5513	TTTGGGCGCGGCTCGGGGTGCAACCCGAATCTCTATCTCTGACTTGTGATCCGGTA
Db	4438	GAATCAGCACTGCAATTTGCTGTCTACAGTGGGTCTGTATATTTACTGAAACTGAT	5515	TTTGGCGGCAAGCTCAGGCTTAATGGCAGAAACCTCTTTCAGTTGACTTTGATGCAACA
Qy	4496	GAGAAAGATCAACCGCATCAGGAGGTCAACCTAGTGGAGAGACTCTACTCGTTGCT	5573	ACACGACACAGAGAAGATCTATGATGACCAACCGCAAGTTCACCTTCGGATTTCTGTACGAC
Db	4498	GAGAAAGATTTAACCGGATAGGCAAGTCAACAGATGGAGAAATCTCTTATGTTGCC	5575	ACAAAGACAGAAAGATCTATGACGACCAACCGTAAATTTCTACTGAGGATCGGCTACGAC
Qy	4556	GGGGCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTTGTTTCTGGAGAC	5633	CAGCGGGGCGGCCAGCTCTGGTCAACCCAGCAGCGCTGATGGTGTCAACGTGACA
Db	4558	GGAAATACCTTCAGAGTGTGACTGTGCAAAATGATGATCCAACTGTGACTGTACAGAGTGA	5635	ACGCTCTGGCACCGGCTCTCTGGCTGCCAAGCAGCAAGCTGATGGCCGCTCAATGTGACC
Qy	4616	GATGTTATGCAAGGATGCAAGTTAAATACCCCATCTCTCTGGCTGTGTGCTGAT	5693	TACTCCCTGGGGTTTACATTTGCTGTCATCCAGGGGCTCATGTCTGAAAGAAATGGAA
			5695	TATTATCCACAGTCAATTTGCCAGATTCACCGAGGACCACTACGAGGAAGTAGAT

QY 5753 TACGACAGGCGGCGGCATCACATCCAGGATCTTCGCTGATGGAGACATGAGCTAC 5812
 Db 5755 TATGACGACAGGAGAGATGCTCTCGGGTCTTGTGATGGTAAACATGAGATTAC 5814
 QY 5813 ACATATCTTAGAGAAGTCCATGGTGTCTACTACACAGCCAGAGGAGTATATCTTTCAG 5872
 Db 5815 ACATATTTAGAAAAGTCCATGGTCTTCTGCTTATAGCCAGCGCAGTACATCTTCGAA 5874
 QY 5873 TTCGACAGATGACCGCTCTCTCTGTGACGATGCCAAAGTGGCGGCGACAGACTA 5932
 Db 5875 TACGATATGTGGGACCGGCTGTCTGCATCACCATGGCCAGTGTGGTTCGCCACACATG 5934
 QY 5933 GAGACCATCCGCTCAGTGGGTACTACAGAAACATCTATCAGCCCGCTGAGGGCAATGCC 5992
 Db 5935 CAGACCATCCGATCCATGGCTACTACCGACATATACAAACCCCGGGAAGCAAGCC 5994
 QY 5993 TCAGTATACAGGACTTCAGTAGGATGGGCACTCTCTTCACTCTTACCTGGGCACT 6052
 Db 5995 TCCATCATCAGGACTACAAAGAGAGGGCTGCTTCTACAAACAGCTTCTTGGGTACA 6054
 QY 6053 GSCCGAGGCTGATATACAAAGTATGGCAAACTGTCAAAGCTGGCAGAGACCTCTATGAC 6112
 Db 6055 AGTCGGAGGCTTATTCAAATACAGAGGAGACTAGGCTCTCAGAAATTTTATATGAT 6114
 QY 6113 ACCACCAAGGTCACTTCCCTATGACGAGACGCGCAGGCTGCTGAAGACCATCAACCTA 6172
 Db 6115 AGCAAGAGTCACTTACCTATGATGAACAGCAGGAGTCTTAAAGACAGTAAACCTC 6174
 QY 6173 CAGATGAGGCTTCACCTGCACCATCCGCTACCGTACAGATTGGCCCTGATTGACCGA 6232
 Db 6175 CAGAGTATGTTTATTTGACCATATAGATACAGGCAAAATGGTCCCTGATTGACAG 6234
 QY 6233 CAGATCTTCCGCTTCACTGAGGAAGGATGGTCAACGCCGTTTGTACTCAACTATGAC 6292
 Db 6235 CAGATTTTCCGCTTTAGTGAAGATGGATGCTAAATGCAAGATTTGACTATAGCTATGAC 6294
 QY 6293 AACAGCTTCCGGGTGACGAGATCGAGCTGTGATCAACGAGACCCCACTGCCATGAT 6352
 Db 6295 AACAGCTTCCAGTGAACGAGATCGAGCTGTGATCAATGAAGCCCACTGCCTATGAT 6354
 QY 6353 CTCTATCGCTATGATGCTCAGGCAAGACAGCAGATTTGGGAAGTTTGGTGTCACT 6412
 Db 6355 CTGTATCAGTTTGTATGATTTCTGGCAAGTTGAGCAGTTTGGAAAGTTTGGAGTTATA 6414
 QY 6413 TACTATGATATTAACAGATCATCAACAGCTGTCTATGACCCACCAAGCAAGCTTTGAT 6472
 Db 6415 TATTATGATATTAACAGATCATTTCTACAGCTGTAATGACCTATACGAAGCACTTTGAT 6474
 QY 6473 GCATATGSCAGATGAAGGAGTACGATATCAGATCTTCCGCTCGCTCATGTACTGATG 6532
 Db 6475 GCTCATGGCCGTATCAAGGAGATTCATATAGATATTCAGGTCGCTCATGTACTGAT 6534
 QY 6533 ACCGTCCAGTATGATAACATGGGCGAGTAGTGAAGAGGAGCTGAAGGTAGGACCTAC 6592
 Db 6535 ACAATTCAGTATGATAACATGGGTCGGGTAAACAGAGAGAGATTAAATATGGGCCCTTT 6594
 QY 6593 GCCAATACCATCCGTACTCTATGATATGATGCTGACGGCGAGCTGCGAGCAGTCTCC 6652
 Db 6595 GCCAACACCAACCAATATGCTTATGAATATGATGTTGATGACAGCTCCAAACAGTTTAC 6654
 QY 6653 ATCAATGACAGCCACTCTGGCGCTACAGCTAGACCTCAATGGGAACCTGCACTTACTG 6712
 Db 6655 CTCAATGAAGAATAATGTGGCGTACAACATGATCTGAATGGAACCTCAATTTACTG 6714
 QY 6713 AGCCTGGGAACAGTGCACGGCTCACACACTACGATGATGACATTCGCGACCCGATCACT 6772
 Db 6715 AACCCAGTAAACAGTGGCGCTGTACACCCCTTTCGCTATGACCTGGGAGACAGATCACT 6774
 QY 6773 CGGCTGGTACGTGCAATACAGATGGATGGCTTCTCTGAGGACGCGGCGGT 6832
 Db 6775 CGACTGGGTGATGTTCAATATCGTTGGATGAAGATGGTTTCTTCTACGTCAAGGGGACG 6834

QY 6833 GATATCTTTAGTCAAACTCAGCTGGCTGTCTCATCAAGGCTTACAAACGGGCTGGCAGC 6892
 Db 6835 GAAATCTTTGAATATAGCTTCAAGGGCTTTTAACTCGAGTTTACAGTAAAGGAGTGGC 6894
 QY 6893 TGGAGTGTGAGTACCGCTTACGATGGCCCTGGGGCGCGCTGTCCAGCAAGAGCAGCCAC 6952
 Db 6895 TGGACAGTATCTACCGTTATGACGCCCTGGGAAGCGCTGTTTCTAGCAAAACAGTCTA 6954
 QY 6953 AGCCACCACTCGAGTTCTTCTATGACAGCTGACCAACCCCAACCAAGGTCAACCACCTG 7012
 Db 6955 GGACAGCACTCGAGTTTATGCTGACTTAATCTTATCCCACTAGGATTACTCATGTC 7014
 QY 7013 TACAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGCAGAGCACCTCTTT 7072
 Db 7015 TACAACCACTCCAGTTTCAAGATTAACCTCCCTGTATTATGATCTCCAGAGACATCTTTT 7074
 QY 7073 GCCATGGAGCTGAGCAGTGGTGTAGTATTTTACATAGCTTGTGACAAATCGGACCCCT 7132
 Db 7075 GCCATGGAAATCAGCAGTGGGATGAATTTCTATATTGATCGGATAACACAGGACACCA 7134
 QY 7133 CTGTCTGCTTTAGTGGAAACAGTTTGTATGATCAAGCAAACTCTGTACACAGCTATGG 7192
 Db 7135 CTGGCTGTCTCAGTAGCAATGGGCTTATGCTGAACAGATTAGTACACTGCTCATGG 7194
 QY 7193 GAGATCTACATGGATACCAACCCCACTTTTCAATCATATAGCTATCATGCTGGTGGCTC 7252
 Db 7195 GAAATCTATTTTGACTCTTAATATTGACTTTCAACTGGTAAATGGATTTTCACTGGTGGCTG 7254
 QY 7253 TATGATCCACTCACCAAGCTTGTCCACATGGCGCGGAGATTATGATGCTGCGCCGA 7312
 Db 7255 TATGACCCACTCACCAAAATTAATCCACTTTGGAGAAAGAGATTATGACATTTTGGCAGGA 7314
 QY 7313 CGCTGGAATAGCCAGACACAGAGCTGTGGAAGCACTTTAGTAGCAGCAACGTCATGCT 7372
 Db 7315 CGGTGGAACACACCTGACATAGAAATCTGSA--AAAGAAATTGGGAAGACCCAGCTCT 7371
 QY 7373 TTTATCTCTATATGTTTCAAAAAACAAACCCCACTCAGCAACTCCAGGACATCAAGTGC 7432
 Db 7372 TTTAACTTTGTATGTTTAGGAATAACAAACCTCAGCAAAAAATCCATGACGTGAAGAT 7431
 QY 7433 TFCATGACAGATGTTTAAACAGCTGGCTCACTTTGGATTCCAGCTTACACAAAGTGA 7492
 Db 7432 TACATCAAGATGTTTAAACAGCTGGCTGTGACATTTGGTTTCCATCTGCAAAAGTCTATT 7491
 QY 7493 CTTGTTATCCAAACACAGACATGATGCCATGGAACCCCTCTACGAGCTCATCCACACA 7552
 Db 7492 CTTGATTTCCCTGTTTCCCAAAATTTGATTTAAACAGAACTTCTTACGA-----A 7539
 QY 7553 CAGATGAACACGAGGAGTGGGCAACAGCAAGTCTATCTCGGGGTACAGTGTGAAGTA 7612
 Db 7540 CTTGTGAAGTCAAGCTGGGATGATATACGCCCATCTTTCGGAGTCCAGCAGCAAGTG 7599
 QY 7613 CAGAAAGCAGCTCAAGGCCCTTTGTCACTTTAGAAACGGTTTGACCACTATSGGTCCACA 7672
 Db 7600 GCGCGGAGGCAAGGCCCTTCTGTGCTGGGAAGATGGCGGAGGTGCAAGTGCAGCGG 7659
 QY 7673 ATCACCAGCTGCCAGAGCTCCAAAGACCAAGAGTTTGCATCCAGGGCTCAGTCTTT 7732
 Db 7660 GCGCGGCGGCG--GGCGCGAGTCTCGGTGTGGTTCGCCACCGGTCAAGTCTGATC 7716
 QY 7733 GCAAGGGGGTCAAGTTTGGCTTTGAAGATGGCGGAGTGAACACAGACATCATCAGTGTG 7792
 Db 7717 GCAAGGGCGTCACTGCTGGCGTCAAGCAGGCGCGGTGAGACCAACGCTCAACATC 7776
 QY 7793 GCCAATGAGGATGGCGAGGGTTGTGCTCATCTTTGAACCAATGCCACTACCTAGAGAAC 7852
 Db 7777 GCCAACGAGGACTGCATCAAGGTGGCGGCGGTGTCTCAACACAGCTTCTTACCTGGGAAC 7856
 QY 7853 CTGCATTTCACTTGTGGGTGGATACCCATTTACTTTGTGAACACAGGACCTTTCAGAA 7912
 Db 7837 CTGCATTTCACTTCGAGGGGCAAGACACGCACTACTTTCATCAGACACCAACGCGCGAG 7896
 QY 7913 GGTGACCTGGCCATCTCTGGGCTCAGTGGGGGGCGGCAACCCCTGGAGAAATGGGGTCAAC 7972

Db 7897 AGCGACCTGGGACGCTGCGGTGACACAGCGCGCGAAGCGCTGGAGACGGCATCAAC 7956
QY 7973 GTCACTGTGTCAGATCAACACAGTACTTAAATGGCAGGACTAGCGCTACACAGATC 8032
Db 7957 GTGACGGTGTGCGCATCCACACAGGTGTGTAAACGCGCAGGACGCGCAGGTTGCGCGACGTG 8016
QY 8033 CAGCTCCAGTACGGGGCACTGTGTTGAACACACGCTACGCGGCAACAGTTGGATGAGGAG 8092
Db 8017 GAGATGCAAGTTGCGCGCTGGCGTGCAGTGTGCTACGCGATGACCTGGACGAGGAG 8076
QY 8093 AAGCACCGGTCTGTGAGCTGGCCCGGACAGAGACCGGTGGCCCAAGCGTGGGCCGCGAG 8152
Db 8077 AAGCGCGCATCTGTGAGCAGCGCGGACGCGCGCGCTGGCCCGCGCTGGCGCGCGAG 8136
QY 8153 CAGCAGACACTGCGGAGGAGGAGGAGGCGCTGGCGCGCTGGACAGAGGGGAGAGCAG 8212
Db 8137 CAGCAGCGCGTGGCGAGCGGAGGAGGCGCGCGCTGGACGAGGCGGAGAGCGG 8196
QY 8213 CAGGTGCTGAGCACAGGCGGGTGCAAGGCTACGACGCGCTTTTCGTGATCTCTGTGCGAG 8272
Db 8197 CAGCTGCTGAGCGCGGCAAGGTGCAGGCTACGACGGTACTACGTACTCTCGGTGAG 8256
QY 8273 CAGTACCCAGACTGTGACAGCGGCGCAACACATCCACTTCATGACAGAGCGGAGATG 8332
Db 8257 CAGTACCCCGAGCTGGCGACAGCGCGCAACACATCCAGTTCCTGGCGAGCGAGATC 8316
QY 8333 GCGCGAGGTGAC 8345
Db 8317 GCGAGGAGTAAC 8329

RESULT 12
AX662357
LOCUS AX662357 8473 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 39 from Patent WO02062999.
ACCESSION AX662357
VERSION AX662357.1 GI:29163218
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Spytek,K.A., Li,L., Wolenc,A.R., Vernet,C.A., Eisen,A., Liu,X.,
Malyanar,U., Shimkets,R.A., Tchernav,V.T., Spaderna,S.K.,
Gorman,L., Kekuda,R., Patturajan,M., Gusev,V., Gangolli,E.A.,
Guo,X., Shenoy,S., Rastelli,L., Casman,S.J., Boldog,F.,
Burgess,C.E., Edinger,S., Ellerman,K., Gunther,B., Smithson,G.,
Millet,I. and Macdougall,J.R.
TITLE Proteins and nucleic acids encoding same
JOURNAL Patent: WO 02062999-A 39 15-AUG-2002;
Curagen Corporation (US)
FEATURES
source
1. 8473
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 41.0%; Score 3422; DB 6; Length 8473;
Best Local Similarity 67.1%; Pred. No. 0;
Matches 5027; Conservative 0; Mismatches 2370; Indels 93; Gaps 9;
QY 880 CGGGCACTTCCTTCAAGCTGG---AGGCACCTCCCGGCTCTTCTGACACCATCACC 936
Db 725 CAGGCATTTCCTATTCAAAACAGGAACAGGATACACGCCACTGTGTAGTACTGCAACCCC 784
QY 937 AGGGTACCCTAGCTCCACACAGTGTACTCTCCCGCCCGACCCCTGCCCGCGAG 996
Db 785 AGGATACAAATGGCACTGCTGTTTATACACCTACTCGGCCACTACTTAGAAA 844
QY 997 CACCTTCGCCCGCGCGCTTTAACTCAAGAGCCCTCCAAAGTACTGTAACTGGAAGTG 1056

Db 845 CACCCTATCAAGAAGTGCTTTAAATTCAAGAAGTCTTCAAGTACTGTAGCTGGAATG 904
QY 1057 CGCAGCCCTAGAGCGCCATCGTCACTCAGCCACTCTGTGTCATCTCCTGCTGGCATCTTTGT 1116
Db 905 CACTGCACTGTGTGCGTAGGGGTCTCGGTGCTCCTGGCACTACTCCTGCTCTATTTAT 964
QY 1117 GGCATGCACTGTGTGGCTTAACTGGCACTGGAGCGGATGGAGGGGACAGTGTATGA 1176
Db 965 AGCAATGCAATCTTTGGGCTCAACTGGCAGCTACAGCAGACTGAAATGACATTTGA 1024
QY 1177 GATACGAGGACACAGCCAGAGTTGGCTGTGCAACCGAGCTCTCCTATACCCCTC 1236
Db 1025 GA-----ATGAAAAGTGAATTCATACCATGCGCAACAAACACTGTCTATTACCTTC 1078
QY 1237 AGGGGCACTGGCTTAGACACCCCTGACAGGAAAGCAAGAAACCAACAGAGGAGAGCC 1296
Db 1079 TGAGAGACA-----ATGGAAAAT 1096
QY 1297 CAGTAGTCTTTCTCCAGAGGACAGTTTATAGATTCTGGAGAAATTTGATGTGGAGGGC 1356
Db 1097 AGGTGATTTACGCAAGAAATAACACCATAGATTCCGAGAACTTGATATTGGCCGAAG 1156
QY 1357 AGCCTCCCAAGAGATTCCTCTCGGCACATTTCTGGAGATCTCAAGTGTTCATAGACATCC 1416
Db 1157 AGCAATTCAGAGATTCCTCCCGGATCTTCTGGAGATCACAGCTCTTCAATTGATCAGCC 1216
QY 1417 TGTGATCTGAAATCAATGTGTCTCTGGAAAGGACGCCCTGGTTGGCAATTTATGGCAG 1476
Db 1217 ACAGTTTCTAAATTCATATCTCTTCAGAAAGATGATTTGGAGTATATGGCCG 1276
QY 1477 AAAAGCCTCCCTCTTACATACACAGTTTCACTTTGGAGCTGTGGATGGCAGGAG 1536
Db 1277 GAAAGCCTTACCGCTTCCCATACTACGTATGACTTCTGGAGCTCTGGATGGCAGCAG 1336
QY 1537 GCTCTAACCCAGGAGCGCGGAGGCTAGAGGGGCCCCCGCCAGCTCTCGGGAACCTGT 1596
Db 1337 GCTGATTGCCAGAGACGCGGAGCCTCTTGAGACGGAGAGCGCGCGGAGCGAG 1396
QY 1597 GCCCCCTCCAGCCATGACAGAGGCTTCATCCAGTATTTGGATTCAGGAATCTGGCACTT 1656
Db 1397 ATCCGTCAGCCTTCATGAGCGCGCTTTATCCAGTACTTGGATCTGGAACTCTGGCATCT 1456
QY 1657 GCGTTTTTCAATGACGGAAGAGAGTCAAGTGTGTTTCTTCTCACCACCTGCCATTGA 1716
Db 1457 GCGTTTTTAAATGATGGGAAAATGACAGAGCAGTGTCTTTAATACCACTGTTATAGA 1516
QY 1717 GTCGTTGATTAACCTGCCAGCAACTGCTATGCAATGTGACTGCTCTCTGGGACCTG 1776
Db 1517 GTCGTGTGGATGTCCCGGAATGCCATGGAATGGAGATGGCTTTCTGGAACTTG 1576
QY 1777 CCACCTCTTCTGTGGTTTCTGGGCCCGACTGTGGCAGAGCCTCTGCCCCGCTGCTG 1836
Db 1577 CCATTGTTTCCAGGATTTCTGGTCCGGATTTGTTCAAGAGCGCTCTCCAGTGTATG 1636
QY 1837 TAGCGAATGCGCAATACATGAAAGGACAGTCTGTGTCACACAGTGGCTGGAAGGCGC 1896
Db 1637 TAGTGGCAACGGGACAGTACTCCAAAGGCGCGCTGCTGTTTTCAGCGGTGGAAGGCGAC 1696
QY 1897 TGAGTCCGATGTGCCCAACCAACAGTGTATCGATGTGGCTGCGCAGCAACCATGGCACCTG 1956
Db 1697 CGAGTGTGATGTGCCGACTACCCAGTGTATTGACCCACAGTGTGGGGTCTGGGATTTG 1756
QY 1957 CATCAGGGCAGCTGCACTGCAACCCCTGGCTACAGGGCGAGAGCTGTGAGGAAGTGA 2016
Db 1757 TATCATGGGCTCTGTGTGTTGCAACTCAGGATACAAAGGAAAAGTTGTAAGAAAGCTGA 1816
QY 2017 CTGATGGAACCCCACTGTTCAGGCGGGGTCTGTGCGTGAAGGCGCAATGCAATTCCTT 2076
Db 1817 CTGTATAGACCTGGGTGTCTTAATCATGTGTGTGTATCCACGGGGAATGTCACTGCA 1876
QY 2077 TGTGGATGGGAGGACCAACTGGAGACCCCGAGGCCACATGCTTAGACAGTGTTC 2136

Db 1877 TCCAGGATGGGAGGTAGCAATTTGTAAATATCTGAAGACCATGTGTCCAGACCAAGTGCTC 1936
Qy 2137 AGGCCACGAACTTCTCCCGGACACCGGCTTTGAGCTGTGACCCCAAGCTGGACTGG 2196
Db 1937 CGGCCACGGAACGTATCTTTCAAGAAAGTGGCTCTCCGACGTTGACCCCTAACCTGAGCTGG 1996
Qy 2197 ACAGACTGTTCTATCCAGATCTGTGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTAGG 2256
Db 1997 CCCGACTGCTCAAGAAATATGTTCTGTGACTGTGGCTCACAAGCGGCTTTGCATGGG 2056
Qy 2257 GGGACCTCGCGCTCGAGAGTGGTGGATGGGGGAGCTGCGACAGCGGGGCTGCCA 2316
Db 2057 GGGAGCTGTCTGCTGTGAAGAGGCTGGACGGGCCCAAGCTGTAAATCAGAGAGCTGCCA 2116
Qy 2317 CCCGCGCTGTCCGAGCATGGGACTGCGCGGACGGCAAGTGGAGTGGAGTGGCCCTGGCTG 2376
Db 2117 CCCCCGCTGTCCGAGACGCGGACTGCAAGATGGCAAGTGGATGGAGTGGAGCGAGGCTG 2176
Qy 2377 GAATGGGAAACACTGACCATCGCTCACTATCTGGATAGGGTAGT----- 2421
Db 2177 GAATGGAGACACTGCACTATCGCTCACTATTTGGATAAGATAGTTAAAGACAAGATAGG 2236
Qy 2422 ---TAAAGAGAGTTCCTCGGCTGTGCAATGGCAAGCGAGATGTACCTTAGACCTGAA 2478
Db 2237 ATATAAGAGAGTGTCTGCTGTGTGCAACAGCAATGGAAATGTACCTTCGACCAAAA 2296
Qy 2479 TGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGCTGGCTGTGACACTTCCATGGA 2538
Db 2297 TGGCGGACATTGTGTGTCGCCAGCTGGATGGAGAGGAGCGGCTGTGACGTAGCCATGGA 2356
Qy 2539 GACTGCTCGCGTGACAGCAAGCAATGATGGAGATGGCGCTGGTGGACTGGACCTC 2598
Db 2357 GACTCTTTGCCACATAGCAAGGCAATGAAGGGGATGGACTCAATTGACTGGCATGCC 2416
Qy 2599 TGACTGCTGCGCTCAGAGCCCTGTGCATATCAACCCGCTGTGCGCTGGCTCCCTAACCC 2658
Db 2417 CGATTGCTGCTACAGAGTTCCTGCCAGATCAGCCCTATTTGTGGGGAGCTGCCGATCC 2476
Qy 2659 TCTGGACATATCCAGGAGACAGAGTCCCTGTGTCTACAGAGAACCTA CACTCTTTCTA 2718
Db 2477 TCAGGACATATTAGCCAAAGCCTTCAATGCCCTTCTCAGCAAGCTGCCAAATCCCTTTTA 2536
Qy 2719 TGACCGCATCAAGTTCCTGCGGAGGAGCAGCAGCAGCATATCCCGGGAGAACCC 2778
Db 2537 TGATCGAATCAGTTTCCCTATAGGATCTGATAGCACCCATGTTATACCTGGAGAAAGTCC 2596
Qy 2779 CTTTGATGGAGGATGCTGTGTTATTCGTGGCCAAAGTATGACATCAGATGAAACCC 2838
Db 2597 TTTCAATAAGAGCCTTGCACTGTCTATCAGAGGCCAAGTACTGACTGCTGATGAACTCC 2656
Qy 2839 CCTGGTTGGTGAACATCAGTTTGTCAATAACCTCTCTTTGGATATACATCAGCAG 2898
Db 2657 ACTTATTGGAGTAATGCTCGTTTTTCCATATCCAGAAATATGGAATATCTATTACCCG 2716
Qy 2899 GCAAGATGGCAGCTTTGACTTTGTGCACAAATGGCGGATCTCCATCATCTCGGGTTGCA 2958
Db 2717 CCAGAGCGAATGTTGACTTTGGTGGCAATGTTGGGGCTCTCTAACTTTGGTATTGA 2776
Qy 2959 CGGGCACCTTTCAATCAGAGGAGCACCCCTGTGCTGCTGCGATCGGATCGTTTGT 3018
Db 2777 ACGATCCCATCTCTCACTAGTATCATCTGTGTGGATTCCATGGAAATGCTTTTATGT 2836
Qy 3019 CATGAAACCATCATCATGACATGAGGAGATGAGATTCACAGCTGTGACCTGAGCAA 3078
Db 2837 GATGATACCTTAGTCATGGAGAAAGAGAGATGACATTCACAGCTGTGATCTGAGTGG 2896
Qy 3079 TTTTGGCGGCCCAACCCAGTGTCTCTCCATCCCACTGACGTCCTTCGCGAGCTCTG 3138
Db 2897 ATTGCTGAGGCGCAATCCCATCATTTGTGTCATCACTTTTATCCACCTTTTTCAGATCTTC 2956
Qy 3139 TGCAGAGAAAGGCCCATTTGCGCGGAAATTCAGGCTTTGACGAGGAAATCTCTATCTC 3198
Db 2957 TCCTGAGAGAGTCCCATCATCTCCGAAAACAGAGTACTCCAGAGGAAATCAATCTCC 3016

Qy 3199 TGGCTGCAAGATGAGGCTGAGCTACTGAGCAGCGGACCCCTGGCTCAAAATCTGTCT 3258
Db 3017 AGGAACAGATTTGAACCTCTCTACTTTGAGTTCCAGAGCTGAGGGTATAGTCAGTTCT 3076
Qy 3259 GAGGATCAGCTTCACCCACCGACCATCCCTTCAACCTCATGAAGTGCACCTCATGGT 3318
Db 3077 CAAGATCACCATGACCCAGTCTATTTATTTCCATTTAATTAATGAAGGTTCTATCTATGGT 3136
Qy 3319 ACGGTGGAGGCGCCCTCTTCAGGAAGTGTTCGTCGACGCCACAGACTGTCTATT 3378
Db 3137 AGCTGTAGTAGGAAGACTCTTCCAAAAGTGGTTTCTGCTCCACAAAATCTTGGCTATAC 3196
Qy 3379 TTTCTATTTGGGCAACAGACAGCTCTACAACAGAGGTTTGGGCTTTTTCAGAGGCTT 3438
Db 3197 TTTCTATTTGGGATAAAACAGATGCATATAATCAGAAAGTCTATGTCTATCTGAAGCTGT 3256
Qy 3439 TGTTCCTGGTGTATGAATATGAATCTCTGCCAGATCTAATCTCTGTGGGAAAAAGAAC 3498
Db 3257 TGTGTGAGTTGGATATGATGATGAGTCTGTGTTGGACCTGACTCTGTGGGAAAAAGAGAC 3316
Qy 3499 AACAGTGTGAGGCTATGAAATTCAGCGCTTCCAAAGCTTGGAGGATGGAGCCTAGACAA 3558
Db 3317 TGCCATTTCTGACAGGCTATGAATTTGGATGCGTCCACATGGTGGCTGACATTAGATAA 3376
Qy 3559 ACATCATGCCCTCAACATTCAAAGTGGTATCTCTGCACAAAGGAAATGGGGAGAACAGT 3618
Db 3377 ACATCAGCTGTGGATGTACAGAACGTTATCTGTACAAAGGAAACCGGGAAAAACCAATT 3436
Qy 3619 TGTGTCTCAGGAGCTCTGTCTATTTGGGAGCATCATGGCAATGGGCCCGGAGAGCAT 3678
Db 3437 CATCTCCAGAGCCTCCAGTCTGATGATGATCATGGGCAATGGGCAAGGCGCAGCAT 3496
Qy 3679 CTCCTGCCCGAGCTCAACCGCTTGTGACGGCAACAAAGCTCTGCGCCCGCAGTGGCCCT 3738
Db 3497 TTCCTGCCCGAGTTGCAATTTGTAAGCTGATGGTAAACAAAGTTACTGGCCCCAGTGGCGCT 3556
Qy 3739 CACCTGTGGCTCTGACGGAGGCTCTATGTTGGTGTATTTCACTACTAGAACATCTT 3798
Db 3557 AGCTTGTGGATTCGATGGCAGTGTGATGAGGAGTTTCACTACGTGCGGCGGATATT 3616
Qy 3799 CCGCTCTGGAATGTCACCAACATCTAGAGCTGAGGAATTAAGATTTTTCAGACATAGTCA 3858
Db 3617 CCGTTCTGGAATGTAAACAAGTCTTAGAACTAAGAAATTAAGATTTTTCAGACATAGCAG 3676
Qy 3859 GAGTCAGACACAAATATCTACCTGGCCACAGACCCCATGAGTGGGCGGCTTCTCTTTC 3918
Db 3677 CAACCCAGCTCATAGATATCTACCTTGCAACGGATCCAGTACGGGAGATCTGTACGTTTC 3736
Qy 3919 TGACAGCAACAGCGCGGGTCTTTTAAATCAAGTCCACTGTGGTGGTGAAGGACCTTGT 3978
Db 3737 TGACCAAAACCCCGCAAAATTTATCGCCCAAGTCACTTACGGGGGCAAAAGACTTGAC 3796
Qy 3979 CAAGAACTCTGAGTGTGTCGGGGACAGGTGAACAGTGCCTCCCTTTTGTATGACATCG 4038
Db 3797 TAAAAATGCAAGTCTGTCGACGGGACAGGGAGCAATGCTTCCGTTTTCAGAGGCGAG 3856
Qy 4039 CTGCGGGATGTGGGAAGGCCACAGAACCCACATCTACCAATCCAGGGGTATTACAGT 4098
Db 3857 ATGTGGGATGGAGGAAGGCCGTGGAAGCCACACTCATGAGTCCCAAGGAATGGGAGT 3916
Qy 4099 GGACAAGTTTGGGCTGATCTACTTCTGTTGGATGGCAACCATGATGACAGCGCATCGATCGAA 4158
Db 3917 TGATAAAGATGGATTAATCTACTTTGTTGATGGAACCATGATAGGAAAGTTGACCAAAA 3976
Qy 4159 TGGATCATCTCCACCTGCTCGGCTCTAATGATCTCACATCAGCCCGGCGCACTCAGCTG 4218
Db 3977 TGGAAATCATCACTCTCTTGGGCTCTAACGATTTGACTTCAGCCAGACCTTTTAACTTG 4036
Qy 4219 TGATTCTGTCTATGGATATTTCCAGGTAAAGTGGGAGTGGGCGCCACAGACTTAGCCATCA 4278
Db 4037 TGACACAGCATGCACATCAGCCAGGTAGCTGTGGAATGGGCCACTGACCTTAGCCATTAA 4096

QY 4279 CCCAATGGACACTCATTATGCTCCTCGAACAACAAATGGTGGTCTCGAAATCTCTGAAA 4338
DB 4097 CCTATGATGATCACTCCATTTATGCTCTGGATAAATAGTAGTTTACAGATCACTGAAA 4156
QY 4339 CCACAGGTGGCCATTTGCGCGGAGGCCCATGCACTGCCAGGTCCCTGGCATTGACCA 4398
DB 4157 TCGTCAAGTTCGCAATGCTGTGGACGCCCCATGCACTGTTCAGGTTCCCGAGTGG--A 4213
QY 4339 CTTCTGTCTAAGCAAGGTGGCCATCCAGCAACCTTGGAGTCAGCCACCGTTTGGCTGT 4458
DB 4214 ATATCCTGTGGGAAGCAGCGGGTGCAGACAACATGGAAATCAGCCCATGCAATGTCTGT 4273
QY 4459 TTCACAAATGGGTCTCTATATGCTGAGACTGATGAGAAAAGATCAACCCCATCAG 4518
DB 4274 GTCCTACAGTGGGTCTCTATCTACATTACGAACTGATGAGAAGAAATTAACCGATAG 4333
QY 4519 GGAGTCAACCTAGTGGAGAGATCTCACTCGTTGCTGGGCCCCCGAGTGGCTGTGACTG 4578
DB 4334 CGAGTCAACAAGATGAGAGAAATCTCTTAGTGGCCGGAATACCTTCAGAGTGTGACTG 4393
QY 4579 TAAAAATGATCCCACTGTGATTTCTGGAGAGATGGTTATGCCAGGATGCAAA 4638
DB 4394 CAATAATGATCCCACTGTGACTGTTACAGAGTGGAGATGGCTACGCCAAGGATGCCAA 4453
QY 4639 GTTAAATACCCCATCTCTCTGGCTGTGTGCTGATGGGAGCTCTACGTGGCCGACCT 4698
DB 4454 ACTCAGTCCCCATCTCTCCCTGGCTCTCTCCAGATGGTACACTGTATATTGAGATCT 4513
QY 4699 TGGGAATCCGAATTCGGTTATCCGGAAGAAACAAGCCCTTCTCAACACCCAGAACAT 4758
DB 4514 AGGGAATATCCGGATCCGGGTGTGTCAAGAATAAGCCCTTACTTAACCTCTATGAACCT 4573
QY 4759 GTATGAGTGTCTTACCAATGACCAAGAGCTCTATCTGTTTGTATACCAACCCGCAAGCA 4818
DB 4574 CTATGAAGTGGGTCTCAACTGATCAAGAACTCTACATCTTTGACATCAATGTACTCA 4633
QY 4819 CCTGTACCCAAAGCTTCCGACAGAGACTACTCTGTAACTTCACTTACCTACCTGGGA 4878
DB 4634 CCAATATACTGTAAGTGTAGTCACTGTGTGTTACCTTTTACAAATTTAGCTACAGCAATGA 4693
QY 4879 CGGGACATCACTCATCAAGACAACAATGGCAACATGTAATGTCGCGCGAGACTC 4938
DB 4694 CAATGATATTAAGTGTGACAGACAGCAATGGCAACCCCTTAGAATTAAGACGGGACCC 4753
QY 4939 TACTGGATGCCCTCTGGCTGGTGGTCCAGATGGCAGGTGACTGGGTGACCATGGG 4998
DB 4754 AAATCGCATCGAGTTCAGTGGTGTCTCTGATAAACCAAGTATGATGTTGACAAATAGG 4813
QY 4999 CACCAACAGTGCATCAGAGTGTGACCAACAGAGACAGAGTTGGCCATGATGACATA 5058
DB 4814 AACAAATGGATGTTGAAAGGCATGACTGCTCAAGGACTGGAATTAGTTTGTGTTACTTA 4873
QY 5059 CCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAAACGATGGACACATTTTA 5118
DB 4874 CCATGGCAATAGTGGCTTTTAGCCATTAAGTGTGAACCTGGATGGACAGCTTTT 4933
QY 5119 TGAGTACAGAGTTTGGCCCTTGACAAATGTGACCTTCCTACTGGCCAGGTGAGCAG 5178
DB 4934 TGAATATGACAGTGAAGTCTCTGACAAATGTTACGTTTCCAACTGGAGTGGTCAAAA 4993
QY 5179 TTTCCGAAGTATACAGACAGTTCAGTGCATGTCAGGTAGAGACTCCAGC--AAGA 5235
DB 4994 CTTGATGGGACATGACAGAGGTATCAAGTGGACATTAAGTCACTTAGCCGAGAGA 5053
QY 5236 TGATGTCAACCAACCAACCTGTCTGCTCGATAGGCGCTTCTACACACTGTGTGCAAGA 5295
DB 5054 AGATGTGAGCATCACTTCAAAATCTGTCTCGATCGATTCCTTCTACACCATGGTTCAAGA 5113
QY 5296 CCAAGTCCGGAACAGCTACTACATCGGGCCGATGGCTCTCTGGGCTGTCTGGGCCAA 5355
DB 5114 TCAGTTAAGAAACAGTACCAAGTTGTTATGACGGCTCCCTCAGAAATTAATCTACGCCAG 5173
QY 5356 CGGCATGGAGGTGGCGGTGCAGACTGAGCCCCCACTTGTGCTGGCTGGCACCCTCAACCCAC 5415

DB 5174 TGCCCTGGACTCACACTCAAAACAGAGCCGACGTTCTGGCTGGCACCGTAATCCGAC 5233
QY 5416 CQTGGCAAGAGAAATGTCACTGCTGCCATCGACACAGCCCTCAACCTGGTGGAGTGGG 5475
DB 5234 GGTTCGCAAAAGAAACATGACTTTGCTGGCGAGAACGCTCAAACTTGGTGAATGGAG 5293
QY 5476 CMAGCCCAAGAGACAGGCTCGGGGCCAGGTCACTGCTTTTGGGCGCGGCTCGGGGTGCA 5535
DB 5294 ATTCCGAAAAGAGCAAGCCCAAGGAAAGTCAATGTCTTTGGCCGCAAGCTCAGGTTAA 5353
QY 5536 CAACCCAAATCTCTATCTCTGACCTTTGATCGGTAAACACACACAGAGAAGATCTATGA 5595
DB 5354 TGGCAGAAACCTCTCTTCACTTGAATCGAACAACAAGACAGAAAAGATCTATGA 5413
QY 5596 TGACCAACGGAATTCACCTTTCGATTCGTACAGAGGGGGGGCCGACGCTCTG 5655
DB 5414 CGACCAACCGTAAATTTCTACTGAGGATCGCTACGACGCTCTGGGACCCGACTCTCTG 5473
QY 5656 GTCAACCCAGCAGAGCTGAATGGTGTCAACGTGTACATCTCCCTGGGGGTTTACATTCG 5715
DB 5474 GCTGCCAAGCAGCAAGCTGATGGCGTCAATGTCACTTATTCATCCACAGGTCAAAATGC 5533
QY 5716 TGGCATCCAGAGGGGCTCATGTCTGAAAAGATGGAATACGACAGGCGGGCGCATCAC 5775
DB 5534 CAGCATCCAGAGGACCACTAGCGAGAAAGTAGATTATACGAGACAGGGAGGATCGT 5593
QY 5776 ATCCAGGATCTTCTGCTGATGGGAAGACATGGAGCTACACATCTTAGAGAAGTCCATGT 5835
DB 5594 GTCTGGGTCTTTGCTGATGGTAAACATGGAGTTACACATTTAGAAAAGTCCATGT 5653
QY 5836 GTGTCTACTACAGCAGAGGAGGAGTATCTTTGAGTTTCGACAAAGATGACCGCTCTC 5895
DB 5654 TCTTCTGCTTCATAGCCAGCGCAGTACATCTTCGAAATACGATATGTGGGACCGCTCTC 5713
QY 5896 TTTCTGAGCAGTCCCAACGCTGGCGGCGAGACATCTAGAGACATCTCGCTCAGTGGGCTA 5955
DB 5714 TGGCATCAGCATGCCAGTGTGGCTGCCACACCATGCAGACCATCCGATCCATTTGGCTA 5773
QY 5956 CTACAGAAACATCTATCAGCCCTTGAGGGAATGCTCAGTACACAGAGCTTCACTCA 6015
DB 5774 CTACCGCAACATATACACCCCGGGAAGCAAGCTCCATCTACCGACTACACCA 5833
QY 6016 GGATGGGCACTCTTTCACACCTTCTACCTGGGCACTGGCCGAGGTTATACAAGTA 6075
DB 5834 GGAAGGCTGCTTCTACAAACAGCTTTCTTGGGTCAAGTCGAGGCTTCTTATCAATA 5893
QY 6076 TGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAAGGTCAGTTTCACTTA 6135
DB 5894 CAGAAGGCAGACTAGGCTCTCAGAAATTTATATGATAGCACAAGAGTCAGTTTACCTTA 5953
QY 6136 TGACGAGCGGAGGCTGCTGAAGACCATCAACCTACAGAATGAGGCTTCACTGAC 6195
DB 5954 TGNAGAAACAGCAGGAGTCTTAAGACAGTAACCTCCAGAGTATGTTTATTTGAC 6013
QY 6196 CATTCGCTACCGTCAAGATTGGGCCCTGATTCGACGACAGATCTTCCGCTTCACTGAGA 6255
DB 6014 CATTAGATACAGGCAAAATGGTCCCTGATTGACAGGAGATTTTCCGCTTTAGTGAAGA 6073
QY 6256 AGGCATGTCACAGCCGCTTTTGAATACAACTATGACACAGCTTCCGGGTGACAGCAT 6315
DB 6074 TGGATGGAATGCAAGATTTGACTATAGCTATGACAAACAGCTTTCGAGTACCAAGCAT 6133
QY 6316 GCAGGTGTGATCAACGAGACCCCATGCCCCATGATCTCTATCGCTATGATGATGTGTC 6375
DB 6134 GCAGGTGTGATCAATGAAACGCCCATGCTTATGATCTGATCAGTTTATGATGACATTC 6193
QY 6376 AGGCAGACAGCAGTGTGGGAAGTGTGGTGTCTATTACTATGACATTAACAGATCAT 6435
DB 6194 TGGAAAAGTGGAGAGTTTGGAAAGTTTGGAGTTATATATTATGATATTAACAGATCAT 6253
QY 6436 CACCACAGCTGTCTATGACCCACCAAGCATTTTTCATGTCATATGCGAGATGAAGAACT 6495

Db 6254 TTCTACAGCTCTAATGACCTATACGAAGCACTTTTGATGCTCATGGCCGTTATCAAGGAGAT 6313
Qy 6496 GCAGTATGAGATCTTCGGCTCGCTCATGTACTGATGACCGTCCAGTATGATTAACATGGG 6555
Db 6314 TCAATATGAGATATTCAGTGGCTCATGTACTGATTAATTCAGTATGATTAACATGGG 6373
Qy 6556 GCGAGTATGAGAGAGAGCTGAAGGTAGGACCTCAGCCAAATACCACTCGCTACTCTTA 6615
Db 6374 TCGGGTACCAAGAGAGAGATTAATAATAGGCCCTTTTCCAAACACCAACCAATATGCTTA 6433
Qy 6616 TGAGTATGATCTGACGCGGAGCTGACAGAGTCTCCATCAATGACAGGCACTCTGGCG 6675
Db 6434 TGAATATGATTTGATGAGAGCTTCCAAACAGTTTACTCAATGAAAAGATATGTTGGCG 6493
Qy 6676 CTACAGCTACGACCTCAATGGGAACCTGCACTTACTAGCCCTGGGAACAGTGCAGGCT 6735
Db 6494 GTACAACTACGATCTGAATGGAAACCTCCATTTACTGAACCCAAAGTAACAGTGGCGTCT 6553
Qy 6736 CACACCACTAGCTATGACATCCGGACCGCATCACTCGGCTGGTGACGTGCAATACAA 6795
Db 6554 GACACCCCTTCGCTATGACCTGCGAGAGAGAACTCACTCGACTGGGTGATGTTCAATCG 6613
Qy 6796 GATGATGAGATGGCTTCTGAGCCAGCGGGCGGTGATATCTTTCAGTACCACTCAGC 6855
Db 6614 GTTGGATGAAGTGTTCCTACGTCAAGAGGCAACGAAATCTTTGAATATAGTCCAA 6673
Qy 6856 TGGCTGCTCATAGGCTCAACCGGGCTGGAGCTGGAGTGTGAGTACCGCTACGA 6915
Db 6674 GGGGCTTCTAATCGAGTTTACAGTAAGGCACTGGCTGACAGTGTACTACCGTTATGA 6733
Qy 6916 TGGCTGGGCGGCGGTTCACAGAGAGAGCAGCAGCCACCTCGAGTCTTCTTA 6975
Db 6734 CGGCTGGGAAGGCGTGTCTTACAGAAAACAGTCTAGGACAGCACCTGAGTTTTTA 6793
Qy 6976 TGCAGACTGACCAACCCCAAGGTCAACCACTGTACAACTCCAGCTGTGAT 7035
Db 6794 TGCTGACTTAACTTATCCACTAGATTAATCATGTCTTCAACACCTTCGAGTTCAGAAAT 6853
Qy 7036 CACTCCCTTACTACGACTTGCAGGACACCTCTTGGCATGAGCTGAGCTAGTGTGA 7095
Db 6854 TACCTCCCTGATTATGATCTCCAGGACATCTTTTGGCAATGGAATCAGCAGTGGGA 6913
Qy 7096 TGAGTTTACATAGCTTGTGCAACATCGGACCCCTCTTGTCTGTTTGTAGTGAACAGG 7155
Db 6914 TGAATTTATATGATCGGATTAACAGAGGACACCACTGGCTGTGTTAGTAGCAATGG 6973
Qy 7156 TTTGATGATCAAGCAATCTGTACAGCCTATGGGAGNCTACATGATACCACTCC 7215
Db 6974 GCTTATGCTGAACAGATTCAGTACCTGCAATATGGGGAATCTATTTGACTTAATAT 7033
Qy 7216 CAACCTTTCAGATCATATAGCTACCACTGAGTGGCTCTATGATCCACTCAACAGCTGT 7275
Db 7034 TGACTTTCACTGTAATGGAATTCATGTTGGCTGTATGACCCACTCCACCAATTAAT 7093
Qy 7276 CCATATGGCGCGGAGATTAATGATGCTGGCGGACGCTGGAGCTAGCCACGACCA 7335
Db 7094 CCACCTTGGGAAGAGATTAATGACATTTTGGCAGGACGCTGGACAAACCTGACATAGA 7153
Qy 7336 GCTGTGGAAGACCTTACTGACGACAGCTCATGCTTTTAACTCTATATGTTCAAAA 7395
Db 7154 AATCTGGA--AAAGATTTGGAGAGCCAGCTCTTTTAACTTTGACATGTTTAGAA 7210
Qy 7396 CAACAACCCCATCAGCAATCCAGGACATCAAGTGTTCATGACAGATGTTAACAGCTG 7455
Db 7211 TAACAACCCCTGCAAGCAAAATCCATGAGCTGAAAGATTATACATCAGATGTTAACAGCTG 7270
Qy 7456 GCTGCTACCTTTGATTCAGCTACCAAGCTGATCCCTGGTTATCCCAACAGACAT 7515
Db 7271 GCTGTGACATTTGGTTTCCATCTGCAAAATGTTATCTCTGGATTCCTGTTCCCAAT 7330
Qy 7516 GGATGCCATGAACCCCTCTACGAGCTCATCCACACAGATGAACAGCAGGAGTGGGA 7575
Db 7331 TGATTTACAGAACCTTCTTACGA-----ACTTGTGAAGAGTCAGCAGTGGGA 7378

Qy 7576 CAACAGCAAGTCTATCTCGGGGTACAGTGTGAAGTACAGAGCAGCTCAAGGCTTTGT 7635
Db 7379 TGATATACGGCCATCTTTCGGAGTCCAGCAGCAAGTGGCGGAGGCCCAAGGCTTCT 7438
Qy 7636 CACCTTAGAACGGTTTGACAGCTCTATGGCTCCACATACCACTGCGCAGCAGGTCC 7695
Db 7439 GTCGCTGGGGAAGATGGCCGAGGTGAGTGAGCGCGCCGCGGCCGCG---GGCGGCA 7495
Qy 7696 AAAGACCAAGAGTTTGATCCAGCGGCTCAGTCTTTTGGCAAGGGGFTCAAGTTTGCTT 7755
Db 7496 GTCCTGGCTGTGTTTCGCCACGCTCAGTCAAGTCTGATCGCAAGGGCGTCTGCTGGCGT 7555
Qy 7756 GAAGATGCGCGAGTACACAGACATCATAGTGTGGCCATAGATGCGGAGAGGT 7815
Db 7556 CAGCCAGGCGCGTGCAGACCAAGCTGCTCAACATGCCCAACGAGGACTGCATCAAGT 7615
Qy 7816 TGCTGCCATCTTGAACCACTGACCTACCTAGAGAACCTGACACTTCACTTTGATGGGT 7875
Db 7616 GCGGCGCTGCTCAACAGCGCTTCTACCTGGAGAACCTGCACCTCACATCGAGGCA 7675
Qy 7876 GATACCCATTAATTTGTAACAGGACCTTCAGAGGTGACCTGGCCATCTTGGGCT 7935
Db 7676 GGCACGCACTACTTTCATCAAGACCAACCGCGAGAGGACCTGGGCACTGCGGT 7735
Qy 7936 CAGTGGGGGGGGAACCCCTGGAGATGGGTCAACGTCACTGTCTCCAGATCAACAC 7995
Db 7736 GACGAGCGCGCAAGGCGCTGGAGACGGCATCAACGTGACGTGTGCGAGTCCACAC 7795
Qy 7996 AGTACTTAATGGCAGACTAGACGTACACAGACATCCAGCTCCAGTACGGGCACTGTG 8055
Db 7796 GGTGTGAACGCGACGCGCGAGTTCGCGACCTGGAGATGAGTTTCGGCGGCTGGC 7855
Qy 8056 CTTGAAACACGCTACGGGCAACGTTGATGAGGAGGAGGACGCGTCTGGAGCTGGC 8115
Db 7856 GCTGACGTGGCTACGGCATGACCTGGAGAGGAGGCGGCTATCTTGGAGCAGGC 7915
Qy 8116 CCGGACAGAGCGGTGCGCCAAAGCTGGGCGCGGACGAGCAGAGACTTGGCGGAGGGA 8175
Db 7916 GCGGACGCGCGCTCGCCCGGCGCTGGCGGAGCAGCAGCGCTGCGCGAGCGGA 7975
Qy 8176 GGAAGCCTTGGCGCTGACAGAGGAGGAGAGCAGAGTGTGACACAGGCGGCT 8235
Db 7976 GGAGGCGCGCGCTCTGGAGGCGGAGAGCGGAGCTGTGAGCGCGGCAAGT 8035
Qy 8236 GCAAGCTACAGCGCTTTTTCGTGATCTGTGACAGTACCCAGAACTGTGACAG 8295
Db 8036 GAGGCGCTACGAGCGGTACTAGTACTCTCGTGGAGCAGTACCCGAGCTGGCGACAG 8095
Qy 8296 CCGCAACCAATCCACTTATGAGACAGGCGAGATGGCGGAGGTGAC 8345
Db 8096 CGCCAAACCAATCCAGTCTTCTGGCGAGGAGATCGCAGGAGGTAAC 8145

RESULT 13

AX662359
LOCUS 8487 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 41 from Patent WO02062999.
ACCESSION AX662359
VERSION AX662359.1 GI:29163219

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

1

Spatel, K.A., Li, L., Wolenc, A.R., Vernet, C.A., Eisen, A., Liu, X.,

Malyankar, U., Shimkets, R.A., Tchernev, V.T., Spaderna, S.K.,

Gorman, L., Kekuda, R., Patturajan, M., Gusev, V., Gangolli, E.A.,

Guo, X., Shenoy, S., Rastelli, L., Casman, S.J., Boldog, F.,

Burgess, C.E., Edinger, S., Ellerman, K., Gunther, E., Smithson, G.,

Millett, I. and Macdougall, J.R.

Proteins and nucleic acids encoding same

TITLE

JOURNAL Patent: WO 02062999-A 41 15-AUG-2002;
Curagen Corporation (US)
FEATURES Location/Qualifiers
source 1..8487
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 40.7%; Score 3398.8; DB 6; Length 8487;
Best Local Similarity 67.0%; Pred. No. 0;
Matches 5008; Conservative 0; Mismatches 2362; Indels 102; Gaps 9;

QY 880 CGGGCACTTCCTCTTCAAGCTGG--AGGCACCTCCCGCTCTCTGACACATACCC 936
Db |||||
QY 766 CAGGATTTCTTATTCATAACAGGAACAGGTACAAAGCCACTGTTCACTACGAAACCC 825
Db |||||
QY 937 AGGTACCACCTAGCTCCAGCACAGTGTACTCTCTCCGCCCGACCCCTGCCCGCAG 996
Db |||||
QY 826 AGGATACACATGGCATCTGGCTCTGTTTATTCACCACCTACTCGGCCACTACCTAGAA 885
Db |||||
QY 997 CACCTTCGCCCGCGGCTTTAACTCAAGAGCCCTCCAAAGTACTGTAACTGGAAGTG 1056
Db |||||
QY 886 CACCCTATCAAGAGTGCTTTAAATCAAGAGTCTTCAAGTACTGTAGCTGGAATG 945
Db |||||
QY 1057 CCGAGCCTTGAGCGGCATCTCATCTCAGCCACTCTGGTCTCTCTGTCGATCTTTGT 1116
Db |||||
QY 946 CACTGCACTGTGCGGTAGGGTCTCGTCTCTGCAATCTCTGTCTTTATTTAT 1005
Db |||||
QY 1117 GGCATGACCTGTTTGGCTTAACTGGACCTGCGAGCCGATGAGGGGCGAGATGATGA 1176
Db |||||
QY 1006 AGCAATGATCTCTTTGGCTCACTGGCAGCTACAGCAGACTGAAATGACATTTGA 1065
Db |||||
QY 1177 GATCAGGAGACACAGCAGCTGAGTGGCTGTGCCAACCCAGCTCTCCCTATACCCCTC 1236
Db |||||
QY 1066 GA-----ATGAAAGTGAATCTGATACCATGCGCAACAACTGTGTCTTACCTTC 1119
Db |||||
QY 1237 AGGGGCACTGGCTTAGAGACCCCTGACAGGAAAGGCAAGAACACAGAAAGAGCC 1296
Db |||||
QY 1120 TGGAGACA-----ATGGAATTT 1137
Db |||||
QY 1297 CAGTAGTTTCTTCCAGAGGACAGTTTCATAGATCTGGAGAAATTTGATGTGGAAGCG 1356
Db |||||
QY 1138 AGTGGATTTACGCAAGAAATAACACCATAGATTCGGGAACTTGATTTGGCCGAAG 1197
Db |||||
QY 1357 AGCCTCCCAAGATTTCTCTGGCAGCTTCTGGAGATCTCAAGTGTTCATAGACCATCC 1416
Db |||||
QY 1198 AGCAATTCAGAGATTTCTCCGGATCTCTGGAGATCAGCTCTTCATGTATCAGCC 1257
Db |||||
QY 1417 TGTGATCTGAAATTCATATGTCTCTGGGAAAGCGACCCCTGTTGGCATTTATGCGAG 1476
Db |||||
QY 1258 ACAGTTTCTTAAATTCATATCTCTCTTCAGAGGATGCATTGATGGAGTATATGCGCG 1317
Db |||||
QY 1477 AAAAGGCTCTCTCTTCCATACACAGTTTGTGAGTCTTGTGAGCTGCTGGATGCGAGAG 1536
Db |||||
QY 1318 GAAAGGCTTACGGCTTCCCATCTCAGTATGATCTGAGGATCTCTGGATGCGAGAG 1377
Db |||||
QY 1537 GCTCCTAACCCAGGAGCGCGAGCTTAGAGGGACCCCGCGCAGTCTCGGGAACTGT 1596
Db |||||
QY 1378 GCTGATGCCAGAGAGCGAGCGCTGCTTGACACGAGAGAGCCGGCGGCGAGGCGAG 1437
Db |||||
QY 1597 GCCCCCTCCAGCCATGAGACAGGCTTCACTCAGTATTTGATTCAGGATCTGGCACTT 1656
Db |||||
QY 1438 ATCCGTGAGCTTCTATGAGCGCGCTTATCCAGTACTTGGATTTCTGGAATCTGGCATCT 1497
Db |||||
QY 1657 GGCTTTTATCAATGACGAAAGGAGTCAAGAGTGGTTTCTCTTCTCACCCTGCGCATTTGA 1716
Db |||||
QY 1498 GGCTTTTATATGATGGAAATTCAGAGCAGGTGCTTTTAAATACCATTTGTTATAGA 1557
Db |||||
QY 1717 GTCGGTGAATTAATCTGCCCGACCACTGCTATGGCAATGGTGAATCTCTGGGACCTG 1776
Db |||||
QY 1558 GTCTGTGGTGAATGTCCCCGAAATTCGATGGAATGGAAATGGAAATGCGTTTCTGAACTTG 1617
Db |||||

QY 1777 CCACTGCTTCTCTGGTTTCTGGGCCCCGACTGTGGCAGAGCCTCTGCCCCCTGTCTG 1836
Db |||||
QY 1618 CCATTGTTTCCAGGATTTCTGGTCCGATTTGTTCAAGAGCCGCTGTCCAGTGTATG 1677
Db |||||
QY 1837 TAGCGAAATGGCCATATAGAAAGCAGATCTTTGCCACAGTGGCTTGGAAAGCGC 1896
Db |||||
QY 1678 TAGTGGCAACGGCAGTACTCCAAAGGCCGCTGCTCTGTGTTTTCAGCGGTGGAGGCGAC 1737
Db |||||
QY 1897 TGAGTCGATGTGCCCAACCAACAGTGTATCGATGTGGCTGCGACCAACCATGGCACCTG 1956
Db |||||
QY 1738 CGAGTGTGATGTGCCGACTACCCAGTGTATTGACCCACAGTGTGGGGTCTGSGGATTG 1797
Db |||||
QY 1957 CATCAAGGACCTGATCTGCAACCTGCTTACAAAGGGGAGAGTGTGAGAAAGTGA 2016
Db |||||
QY 1798 TATCATGGGCTCTGTGCTGTCAACTCAGGATACAAAGAGAAAGTGTGAAGAACTGA 1857
Db |||||
QY 2017 CTCGATGGACCCCAACATGTTACGGCCGGGTGCTGCTGAGAGGCAATGCCATTGCTT 2076
Db |||||
QY 1858 CTGTATAGACCTTGGGTGTTCTAATCATGTTGTGTATCCACGGGAATGTCACTGCGAG 1917
Db |||||
QY 2077 TGTGGATGGGAGGACCAACCTGCGAGACCCCAAGGGCCACATGTTTAGACAGTGTTC 2136
Db |||||
QY 1918 TCCAGGATGGGAGGTAGCAATTTGAAATTAAGAGACATGTTGCCAGACCAAGTGTCT 1977
Db |||||
QY 2137 AGCCACCGAACTTCTCCCGGACACCGGCTTTGACGCTGTGACCCCAAGCTGGAAGT 2196
Db |||||
QY 1978 CGGCCAGGAACTATCTTCAAGAAAGTGGCTCTGACGTGTGACCTTAACCTGGAAGT 2037
Db |||||
QY 2197 ACAGCACTGTTTATCGAGATCTGCTCCGACTGTGTGGCCATGTGGGTGTGCGTAGG 2256
Db |||||
QY 2038 CCCAGACTGCTCAACGAAATATGTTCTGTGGACTGTGCTCACACGGCGTTTTCATGGG 2097
Db |||||
QY 2257 GGGCACTCCGCTCGAGGATGGCTGGATGGGGGAGCCTGGGACAGCGGGCTTGCCA 2316
Db |||||
QY 2098 GGGGAGCTGTGCTGTGAAGAGGCTGGACGGGCCCACTGTAATCAGAGAGCCTTGCCA 2157
Db |||||
QY 2317 CCGCGCTGTGCGGACATGGGACCTCGCGGACGCGAAAGTGGAGTGCAGGCTCGGCTG 2376
Db |||||
QY 2158 CCCCCGCTGTGCGGACACGGGACCTGCAAGGATGGCAAGTGTGAATGCAGCCATGGCTG 2217
Db |||||
QY 2377 GAATGGCAACACTGCACCATGCTCACTATCTGGAATAGGTTAGTTAAAGAGGTTGCCC 2436
Db |||||
QY 2218 GAATGGAGAGCACTGCATATC-----GAGGGTTGTGCC 2250
Db |||||
QY 2437 TGGGTTGTGCAATGGCAACGGCAGATGTACCTTTAGACCTGAAATGTTGGCACTGCGCTG 2496
Db |||||
QY 2251 TGTGCTGTGCAACAGCAATGGAGATGTACCTTGACCAAAATGCTGTCATTTGTGTG 2310
Db |||||
QY 2497 CCAGCTGGCTGGAGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCTCGGGTGAACAG 2556
Db |||||
QY 2311 CCAGCCTGGATGGAGAGGAGCAGGCTGTGACGTAGCATGGAGACTCTTTTGCACAGATAG 2370
Db |||||
QY 2557 CAAGACATGATGGAGTGGGCTGGTGGACTGCAATGGACCTGACTGCTGCTCCAGCC 2616
Db |||||
QY 2371 CAAGGCAATGAAGGAGATGGACTCATTTGACTGCAATGCGGATTCGCTTACAGAG 2430
Db |||||
QY 2617 CTTGTGCCATATCAACCCGCTGTGCTTGGCTCCCTTAACCTCTGGACATCATCCAGA 2676
Db |||||
QY 2431 TTCTCTGCCAGATCAGCCCTATTGTCGGGACTGCGCGGATCTCAGGACATCATTAGCCA 2490
Db |||||
QY 2677 GACACAGTCCCTGTGTACAGAGAAACCTACCTCTCTTATGACCGCATCAAGTTCT 2736
Db |||||
QY 2491 AAGCCTTCAATCGCCTTCTCAGCAAGCTGCCAAATCTCTTTATGATCAATCAGTTCT 2550
Db |||||
QY 2737 CGTGGCAGGGACAGCAGCAGATATCCCGGGGAGAACCCCTTTGATGGAGGCGATGC 2796
Db |||||
QY 2551 TATAGATCTGATAGCAACCATGTTATACCTGGAGAAAGTCTTTCAATAAGAGCCTTGC 2610
Db |||||
QY 2797 TTGTGTTATTCGTGGCCAAAGTATGACATCAGATGAAACCCCTCTGTTGTGTGAACAT 2856
Db |||||
QY 2611 ATCTGTCACTCAGAGGCCAAGTACTGACTGCTGATGGAATCCCACTTATTTGAGTAAATGT 2670
Db |||||
QY 2857 CAGTTTGTCAATAACCCCTCTCTTTGGATATACATCAGAGGCAAGATGGCAGCTTTGA 2916
Db |||||

Db 2671 CTCGTTTTTCCATTACCCAGAAATATGATATCTATTACCGCCAGGACGGAATGTTGA 2730
 QY 2917 CTTGGTGACAAATCGCGGCATCTCCATCATCCTCGGTTTCGAGGGGACCTTTTCATCAC 2976
 Db 2731 CTTGGTGCAAAATGGTGGGCTCTCTAACTTTGGTATTTGAAGCATCCCAATTCCTCAC 2790
 QY 2977 ACAGAGCACACCCTGTGGCTGCCATGGGATCGCTTTCTTGTGTCATGGAACCAATCATCAT 3036
 Db 2791 TCAGTATCATACTGTGTGATTCATGGAATGCTTTTATGTGATGATACCCCTAGTCAT 2850
 QY 3037 GAGCATGAGAGAAATGAGATCCAGTCCAGTGTGACTGAGCAATTTTCCCGCCGCCAACCC 3096
 Db 2851 GGAGAAAGAGAGAAATGACATTCACGCTGTGATCTCGAGTGGATTCGTGAGGCCAAATCC 2910
 QY 3097 AGTCGTCTCTCCATCCCACTGAGTCTTCGCGAGCTCCTGTGACAGAAAGGCCCAT 3156
 Db 2911 CATCATTTGTCTATCACCTTTATCCACTTTTCAGATCTTCCTGAAGCAGTCCCAT 2970
 QY 3157 TGTCCGGAATTCAGGCTTTTCAGGAGGAATCTCTATCTCTGGCTGCAAGATGAGGCT 3216
 Db 2971 CATTCGCCAAACACAGGTAATCCAGGAACTACAAATTCAGGAAACAGATTTGAACT 3030
 QY 3217 GAGTACCTGAGCAGCGGACCTCGGTACAAATCTGTCTGAGGATCAGCTTCACCCA 3276
 Db 3031 CTCCTACTTGATTCAGAGCTGCAGGCTATGATCAGTCTCAAGATCACCATGACCCA 3090
 QY 3277 CCCGACCATCCCTTCAACCTCATGAAGGTGCACCTCATGTAGCGGTGAGGCGCCCT 3336
 Db 3091 GTCATTTATTCATTTAATTAAGAGTTTCATCTTATGTTAGTCTGTAGTGAAGACT 3150
 QY 3337 CTTGAGGAAGTGTTCGCTGAGCCCGACCTGCTCTATTTATTTCTATTTTGGACAGAC 3396
 Db 3151 CTTCCAAAAGTGTTCCTGCTCACCACAACTTGGCTTATCTATCTATCTGATGGGATAAAC 3210
 QY 3397 AGAGCTCTACAAACAGAGGTGTTCGGCTTTTCAGAGGCTTTGTTTCCTGGGTTATGA 3456
 Db 3211 AGATGATATATCAGAAAGTCTATGTTCTATCTGAGTCTTGTGTCAGTTGATGA 3270
 QY 3457 ATATGAATCTCTGCCAGATCTAATCTCTGTGGGAAAGAAACAAACAGTGTGTCAGGCTA 3516
 Db 3271 GTATGAGTCGTGTTGGACCTGACTCTGTGGGAAAGAGGACTGCCATCTGACGGCTA 3330
 QY 3517 TGAAATTCAGCGCTCCAGCTTGAGGATGAGCTAGACAAACATCATGCCCTCAACAT 3576
 Db 3331 TGAATTTGATGCTCCACATGTTGGTGTGCTGGACATTAGATTAACATCACGTCTGGATG 3390
 QY 3577 TCAAAAGTGTATCTCTGCAAAAGGAATGGGGAGAACAGTTTGTCTCAGCAGCTCC 3636
 Db 3391 ACAGAAAGGTATCTGTTCAAGGAAACCGGGAAACCCAGTTCTATCTCCAGCAGCTCC 3450
 QY 3637 TGTCAATGGAGCATCATGGCAATGGGCGCGGAGAGCATCTCCTGCCCCAGCTGCAA 3696
 Db 3451 AGTCGTGATGATCATATGGCAATGGGCGAAGGCGAGCATTTCTGCCCCAGTTGCAA 3510
 QY 3697 CGGCTTGTGACGGCAACAAGCTCTCTGCGCCCAAGTGGCCCTCAGCTGTGCTCTGACGG 3756
 Db 3511 TGGTCAAGCTGATGTTAACAAGTTACTTGGCCCAAGTGGCGCTAGCTTGTGGATCGATGG 3570
 QY 3757 GAGCTCTATGTTGGTGTATTTCACTACATTTAGAAAGGATCTTCCTCTGGAATGTAC 3816
 Db 3571 CAGTCTGTAGTGGAGTATTTCACTAGCTGCGGGGATTTTCCCTTCTGGAATGTAAAC 3630
 QY 3817 CAACATCTTAGAGCTGAGGAATTAAGATTTTCAGACATAGTCACAGTCAGACACAATA 3876
 Db 3631 AGTGTCTTTAGAACTTAAGAAATTAAGATTTTAGACATAGAGCAACCCAGCTCATAGATA 3690
 QY 3877 CTACTGGCCACAGACCCCATGAGTGGGCGCTCTTCTCTTCTGACACACACCGCGCG 3936
 Db 3691 CTACCTTGCACCGGATCCAGTCACGGGAGATCTGTACGCTTTCTGACACAAACACCGCAG 3750
 QY 3937 GGTCTTTAAATCAAGTCCACTGTGCTGGTGAAGACCTTGTCAAGAACTCTGAGGTGGT 3996

Db 3751 AATTTATCGCCCAAGTCACCTTACGGGGCAAAAGACTTGACTAAAAATGCAGAGTCGT 3810
 QY 3997 TCGGGGACAGGTGACCAAGTCCCTCCCTTTTATGATGACACTCGCTGCGGGATGTTGGAA 4056
 Db 3811 CGCAGGGACAGGGAGCAATGCCCTTCGTTTACGAGCGAGATGTTGGGATGAGGGAA 3870
 QY 4057 GSCCAGACAGCCACACTCAACAATCCAGGGTATTACAGTGACAGAGTTTGGGCTGAT 4116
 Db 3871 GSCCGTGAAGCCACACTCATGAGTCCCAAGGAATGSCAGTTGATGAATGAATTAAT 3930
 QY 4117 CTACTTCGTGGATGGCCACCATGATCAGACGCATCGATCAGAAATGGGATCATCTCAACCT 4176
 Db 3931 CTACTTTTGTGATGAACCAATGATGAGAAAGTTGACCAAAATGGAATCATCACTCT 3990
 QY 4177 GTCGCGCTTAATGATCTCATCATGACCCGGGCACTGAGCTGTGATTTCTGTATGGATAT 4236
 Db 3991 TCTGGGCTTAACGATTTGACTTCAGCCAGACCTTTAACTTTGTGACACACGATGCACAT 4050
 QY 4237 TTCCAGGTGAAGCTGGAGTGGCCACAGACTTAGCCATCAACCAATGGAACACTCACT 4296
 Db 4051 CAGCCAGGTGATCTCTGGAATGGCCCACTGACTAGCCATTAACCTATGATTAATCCAT 4110
 QY 4297 TTATGTCTCGACAAACAATGTGTCCTGCAAAATCTTGAAACACCAAGGTGCGGATGT 4356
 Db 4111 TTATGTCTGGATAAATAGTTTACAGATCACTGAAATCGTCAAGTTCGATTCG 4170
 QY 4357 CGCGGAGGCCATGCACTGCCAGTCCCTGGCATTCACCACTTCCTGTGAAGCAAGGT 4416
 Db 4171 TGCTGGAGCGCCATGCACTGTGAGTTCCCGGAGTGG---AATATCTGTGGGAAGCA 4227
 QY 4417 GSCCATCCACGCAACCTCTGAGTTCAGCCACCGCTTTGGCTGTTCACAAATGGGGTCT 4476
 Db 4228 CGCGGTGCAGACAAACACTTGAATCAGCCACTGCCATTCCTGTGCTCCTACAGTGGGCTCT 4287
 QY 4477 GTATATGCTGAGACTGATGAGAAAGATCAACCGCATCAGGAGGTCAACCATAGTGG 4536
 Db 4288 GTACATTAAGTAACTGATGAGAAAGATTAACCGGATAAGGAGGTCAACACAGATGG 4347
 QY 4537 AGAGATCTCACTCGTTGTGGGGCCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTG 4596
 Db 4348 AGAAATCTCTTAGTGGCGGAAATACCTTCAGAGTGTGACTGCAAAATGATGCCAACTG 4407
 QY 4597 TGATTTGTTTTCTGAGACGATGGTTATGCCAAGGATGCAAGTTAAATACCCATCTTC 4656
 Db 4408 TGACTGTTACACAGATGGAGATGGCTACGCCAAGGATGCCAACTCAGTGGCCCACTCTC 4467
 QY 4657 CTTGGCTGTGCTGTGATGGGAGCTCTACGTGCCGACCTTGGGAAACATCCGAATTCG 4716
 Db 4468 CTTGGCTGTTCTTCAGATGGTACACTGATATTCAGATCTAGGGAATATCCGATCCG 4527
 QY 4717 GTTTATCCGGAAGAAACAGCTTTCTCAACACCCAGAACATGATGAGCTGTCTTACC 4776
 Db 4528 GCGTGTGTCAAAGATAAGCCCTTTACTTAACTCTATGAATCTCTATGAATGCTCTCC 4587
 QY 4777 AATTGACAGGAGCTCTATCTGTTGATACACCGGACGACCTGTACACCAAGCT 4836
 Db 4588 AACTGATCAGAACTCTACATCTTTGACATCAATGGTACTCACCATATCTGTAACTT 4647
 QY 4837 GCCCACAGGAGACTACCTGTACAACCTTCACTTACACTGGGACGCGGACATCACAATCAT 4896
 Db 4648 AGTCATGTTGATTTACTTTTACAATTTTAGCTACAGCAATGACATGATATTTACTGCTGT 4707
 QY 4897 CACAGACAAACAATGGCAACATGGTAAATGTCGCGGAGACTCTACTGGGATGCCCTCTG 4956
 Db 4708 GACAGACAGCAATGGCAACACCCCTTTAGAAATTAGACGGGACCCCAATCGCATGCCAGTTCG 4767
 QY 4957 GCTGGTGGTCCAGATGGCCAGGTGTACTGCGGTGACCATGGGACCAACAGTGCACACTCAA 5016
 Db 4768 AGTGGTGTCTCTGATTAACCAAGTGTATGTTGACATAGGAACAAATGGATGTTGAA 4827
 QY 5017 GAGTGTGACACACAAGGACAGAGTGGCCATGATGATACCATGGAATTCGGGCT 5076
 Db 4828 AGGCATGACTGCTCAAGGACTGGAATTTGTTTGTACTTACCATGGCAATAGTGGCT 4887

QY 5077 TCTGCAACCAAAAGCAATGAAACGGATGGAACAACATTTTATGAGTACGACAGCTTTGG 5136
Db 4888 TTTAGCCCACTAAAAGTATGAAACTGGATGGACAAACGTTTTTGTGACTATGACAGTGAAGG 4947
QY 5137 CCGCTGTGACAAATGTGACCTTCCCTTACCTGGCCAGGTGAGCAGTTTCCGAAAGTGATACAGA 5196
Db 4948 TCGTCTGACAAATGTGACCTTCCCACTGGAGTGGTACAAACCTGATGGGACATGGA 5007
QY 5197 CAGTTCACTGATGCTCAGGTAGAGACCTCCAGC---AAGATGATGTACACATAAACAC 5253
Db 5008 CAAGCTATACAGTGGACATGAGTCACTAGCCGAGAAAGATGTCAAGCATCACTTC 5067
QY 5254 CAACCTGTCTCCTCAGCGCCTTCTACACACTGTCGACACCAAGTCCGGRACAGCTA 5313
Db 5068 AAATCTGTCTCGATCGATTCTTTCTACACCATGTTCAAGATCAGTTTAAAGAAACAGCTA 5127
QY 5314 CTACATCGGGCCGATGCTCCTTTCGGCTCTCTGCGCCAAACGGCATGGAGTGGCGCT 5373
Db 5128 CCAGATTGGTTATGACGGCTCCCTCAGAAATTTATAGCCAGTGGCCCTGGACTCACACTA 5187
QY 5374 GCAGACTGAGCCCACTTGTCTGGCTGGCGACCGCTCAACCCACCGTGGCAGAGGAATGT 5433
Db 5188 CCAACAGAGCCGACGTTCTTGGCTGGCACCGCTAATCCGACGTTGCCAAAAGAAACAT 5247
QY 5434 CACGCTGCCATCGACAAACGGCCTCAACCTGGTGGAGTGGCGCCAGCGCAAGAGCAGGC 5493
Db 5248 GACTTTGGCTGGCGAGAACGGTCAAACTTGGTGGATGGAGATTCGAAAGAGCAGC 5307
QY 5494 TCGGGCCAGTCACTGTCTTGGCGCGCGCTCGGGGTGCACAAACGAAATCTCCTATC 5553
Db 5308 CCAAGGGAAGTCAATGTCTTGGCGCGCAAGCTCAGGGTTAATGGCAGAAACCTCCTTTC 5367
QY 5554 TCTGACTTTCATCGCTPAACAGCACAGAGAAGATCTATGATCACACACCGCAAGTTCAC 5613
Db 5368 AGTTGACTTTGATCGAACACAAAGACAGAAAGATCTATGACACACCGTAAATTTCT 5427
QY 5614 CTTTGGATTCGTACGACAGCGGGCGGCCAGCCTCTGGTCAACCAGCAGCAGGCT 5673
Db 5428 ACTGAGGATCGCTACGACAGCTCTGGCACCGGACCTCTCTGGCTGCGAAGCAAGCT 5487
QY 5674 GAATGGTGTCAAGTGCATACCTCCCTGGGGGTTCATCTGCTGGCATCCAGGGGCAT 5733
Db 5488 GATGCGCTCAATGTCACTTATTCATCCAGAGTCAAATTCAGCATCCAGGAGGCAC 5547
QY 5734 CATGTCTGAAAGATGGAATACGACAGCGCGCGCATCACATCCAGGATCTTCGCTGA 5793
Db 5548 CACTAGCGAGAAAGTAGATTATGACGGACAGGGAGGATCGTGTCTCGGGTCTTTGCTGA 5607
QY 5794 TGGGAAGCATGGAGCTACATACATCTTAGAAGTCCATGGTGTGCTACTACACAGCCA 5853
Db 5608 TGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGTTCTTCTGCTTCATAGCCA 5667
QY 5854 GAGGCAATATCTTTGAGTTGCAAGAAATGACCGCTCTCTCTGTGACGATGCCAA 5913
Db 5668 CGGCAATACATCTTCGATACGATATGTGGGACCGCTGTCTGCCATCACCATGCCAG 5727
QY 5914 CGTGGCGGAGACATCTAGAGACCATCCGCTCAGTGGGTACTACAGAAACATCTATCA 5973
Db 5728 TGTGGCTCGCACACCATGACAGCAATCCGATCCATTTGGCTACTACTACCGCAATATACAA 5787
QY 5974 GCCCCTGAGGCAATCCCTCAGTACATACAGGATTCCTGAGGATGGGACCTCTTCA 6033
Db 5788 CCCCCGGAAGCAAGCGCTCCATCATACGGACTCAAGAGGAGGGGTGCTTCTACA 5847
QY 6034 CACCTTCTACCTGGGCACTGGCCGAGGGTGATATACAAATATGGCAAACTGTCAAAAGCT 6093
Db 5848 AACAGCTTCTTGGTCAAGTCCGAGGGTCTTATTCAAATACAGAAAGGAGCATAGGCT 5907
QY 6094 GCGAGAGCGCTCTATGACACCAAGAGTCAAGTCAAGTTCCCTATGACAGAGCGGAGCAT 6153
Db 5908 CTCAGAAATTTTATATGATAGCAAGAAGTCAAGTCAAGTTTACCTTATGATGAAAACAGCAGGAGT 5967

QY 6154 GCTGAAGACCATCAACCTACAGATGAGGGCTTCCCTGCAACCATCCGCTACCGTCAGAT 6213
Db 5968 CCTAAAGACAGTAAACCTCCAGAGTGTGTTTATTTGCAACATTAGATACAGGCAAT 6027
QY 6214 TGGGCCCCCTGATTGACCCGACAGATCTTCCGCTTCACTGAGGAAGGATGGTCAACCCCG 6273
Db 6028 TGGTCCCCCTGATTGACAGGACAGATTTTCCGCTTTAGTGAAGATGGATGGTAAATGCAAG 6087
QY 6274 TTTTGTACTACATATGACAAACAGCTTCCGGGTGACACAGATGCGAGGTGTGATCAACGA 6333
Db 6088 ATTTGTACTATAGTATGACAAACAGCTTTCCAGTGCACAGATGCGAGGTGTGATCAATGA 6147
QY 6334 GACCCCACTGCCCATTTGATCTCTATCGCTATGATGATGTGTGAGGCAACACAGAGCAGTT 6393
Db 6148 AAGCCCACTGCCCATTTGATCTGTATCAGTTTGTATGACATTTCTGGCAAGATTTGAGCAGTT 6207
QY 6394 TGGGAAGTTTGGTGTCAATTTACTATGACATTAACAGATCATCAACACAGCTGTCTATGAC 6453
Db 6208 TGGAAAGTTTGGAGTTATATATATGATATTAACAGATCAITTTCTACAGCTGTAATGAC 6267
QY 6454 CCACACCAAGCATTTTGTATGATATGCGAGGATGAGGAAGTGCAGTATGAGATCTTCCG 6513
Db 6268 CTATAGAAAGCACTTTTGTATGCTCATGGCCGTATCAAGGAGATTCAATATGAGATATTGAG 6327
QY 6514 CTGCTCATGTATGATGACCGCTCCAGTATGATTAACATGGGGCGAGTAGTGAAGAAGA 6573
Db 6328 GTGCTCATGTACTGGATTACAATTCAGTATGATTAACATGGGTCCGGTAAACCAAGAGAGA 6387
QY 6574 GCTGAAGGTAGACCTCAGCCAAATACACTCGCTACTCTATGATGATGATGCTGACGG 6633
Db 6388 GATTAATAATAGGGCCCTTTTSCCAACACCAACCAATATGCTTATGATATGATGTTGATGG 6447
QY 6634 CCAGCTGCAGACAGTCTCCATCAATGACAAAGCACTCTGGCGCTACAGCTACAGCTCAA 6693
Db 6448 ACAGCTCCAAACAGTTTACTCAATGAAAAGATAATGTGGCGGTACAACTACATCTGAA 6507
QY 6694 TGGGAACCTGCACTTACTGAGCCCTGGGAAACAGTGCACGGCTCACACCACTACGTPATGA 6753
Db 6508 TGGAAACCTCCATTTACTGAAACCAAGTAAACAGTGCCTCTGACACCCCTTCGCTATGA 6567
QY 6754 CATCCGCAACCCATCCTCGGCTGGTGAACCTGCAATACAAAGATGAGTGAAGATGGCTT 6813
Db 6568 CTTGCAAGACAGAACTCACTCGACTGGTGTATGTTCAATATCGGTTGGATGAAGATGGTTT 6627
QY 6814 CTTGAGCGAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCTCTCATCAAGGC 6873
Db 6628 CTTAAGCTCAAAGGGGCAACGGAATCTTTGAAATATAGCTCCAAGGGCTTCTTAACTCAGT 6687
QY 6874 CTACAAACCGGCTGGCAGCTGAGTGTAGGTACCGTACGATGCGCTGGGGGGGGCGGT 6933
Db 6688 TTACAGTAAAGGCACTGGCTGGACAGTATCTACCGTTATGACGGCTGGGAGGCGTGT 6747
QY 6934 GTCCAGCAAGAGCAGCCACAGCCACCTCGAGTCTTCTATGAGACCTCACCACCC 6993
Db 6748 TTCTAGCAAAACCAAGCTTAGGACAGCACTCGAGTTTTTTATGCTGACTTAACCTATATCC 6807
QY 6994 CACCAAGGTCAACCACTGTACAAACACTCCAGCTGAGATCACTCCCTCTACTACGA 7053
Db 6808 CACTAGGATTTACTATGTCTCAACCAATTCGAGTTTCAAGAAATTAACCTCCCTGTATTATGA 6867
QY 7054 CTTGCAAGACACCTCTTTGCCATGGAGCTGAGGATGGTGTGATGATTTTACATAGCTTG 7113
Db 6868 TCTTCAAGGACATCTTTTGGCCATGGAATCAGAGTGGGATGAATTTCTATTGTCATC 6927
QY 7114 TGACAAACATCGGAGACCCCTCTTGTCTTTTGTAGTGAACAGGTTTTGATCAAGCAAAAT 7173
Db 6928 GGATAACACAGGACACCACTCGCTGTGTTTCAAGTGAATGGCTTATGCTCAACACAGAT 6987
QY 7174 CTTGTACACAGCTATGGGAGATCTACATGATACCAACCCCACTTTCAGATCATCAT 7233
Db 6988 TCAGTACACTGATATGGGAAATCTATTTTCACTTAATATTGATTTCAACTGGTAAAT 7047
QY 7234 AGGCTACCATGTGGCTCTATGATCCACTCACCAAGCTTGTCCACATGGGCGCGAGA 7293

Db 7048 TGGATTTTCATGGTGGCTGTATGACCCACTCACCAATTAATCCACTTTGGAGAGAGAGA 7107
Qy 7294 TTATGATGTGTGCGGACCTGGAGCTAGCCAGACCCAGAGCTGTGGAGACCTTAG 7353
Db 7108 TTATGACATTTTGGCAGGACGGTGGACCAACCTGATAGATAATCTGGA---AAGAAT 7164
Qy 7354 TAGCAGCAACCTCATCGCTTTTAATCTCTATATGTTCAAAAACAAACACCCCATCAGCAA 7413
Db 7165 TGGAGAGACCCAGCTCTTTTAATCTGTATGATGTTAGGAATAACAACCTGCAAGCAA 7224
Qy 7414 CTCCAGAGCATCAAGTCTTCATGACAGATGTTTAAAGCTGGCTGTCTACCTTTGAATT 7473
Db 7225 AATCCATGACGTGAAGATTACATACAGATGTTTAAAGCTGGCTGTCTACCTTTGTTT 7284
Qy 7474 CCAGTACACAACTGATCCCTGTTTATCCAAACACAGACATGATGCCATGAACCTC 7533
Db 7285 CCATCTGACAAATGCTATCTCTGATTCCTGTTCCCAATTTGATTTAAGACACCTTC 7344
Qy 7534 CTACAGCTATCCACACACAGATGAATAACAGAGATGGGACAAACAGCAAGTCTATCCT 7593
Db 7345 TTACGA-----ACTGTGAAGATCAGCAGTGGGATGATATACCGCCCATCTT 7392
Qy 7594 CGGGTACAGTGTGAGTACAGAGACGCTCAAGGCTTTGTCACCTTAGAACGTTGA 7653
Db 7393 CGGATCCAGCAGCAAGTGGCGCGCAGGCGCAAGGCTTCTGTGCTGGGAAGATGGC 7452
Qy 7654 CCAGCTCTATGGCTCCCAATCACCAGCTGCCAGAGCTCCAAAGACCAAGAGTTTC 7713
Db 7453 CGAGTGCAGTGAAGCGCGCGCGCGCGG- - -GGCGCGCAGTCTGCTGTGTTTCG 7509
Qy 7714 ATCCAGCGCTCAGTCTTTTGGCAAGGGGTCAAGTTTGCCTTGAAGATGCCGAGTGAC 7773
Db 7510 CACGCTCAAGTCTGATCGCAGAGGGCGTATCTGCGCGTCAAGCAGGCGCGCGTCA 7569
Qy 7774 CACAGACATCATAGTGTGGCAATGAGATGGCGAAGGGTGTCTGCCATCTTGAACCA 7833
Db 7570 GACCAAGTGTCAACATCGCACAGGAGTGCATCAAGTGGCGCGCTGTCTAACAA 7629
Qy 7834 TGCCACATCTAGAGAACCTTGCACTTCAACATTTGATGGGTGGATACCCATTTACTTGT 7893
Db 7630 CGCTTCTACCTGGAGAACCTTGCACTTCAACATCGAGGGCAAGGACACGCACTACTTCAT 7689
Qy 7894 GAACACGACCTTCAGAGTGTACCTGGCCATCTGGGCTCAGTGGGGGGCGGCGAAC 7953
Db 7690 CAAGACACACGCGCGAGAGACCTGGGACACGCTGGGTTGACAGCGCGCGCAAGGC 7749
Qy 7954 CTTGAGAAATGGGTCAACGTCACCTGTGTCCAGATCAACACAGTACTTAAATGCGAGGAC 8013
Db 7750 GCTGGAGAACGGCATCAACGTGACGGTGTGCGAGTCCACACAGGTGGTGAACGCGAGAC 7809
Qy 8014 TAGAGCTACACAGACATCCAGTCCAGTACGGGGCACTGTGTTGAACACAGCTACGG 8073
Db 7810 GCGCAGGTTCCGCGAGTGGAGATGCGATTCGGCGCGCTGGCGTGCACGTGGCGCTACGG 7869
Qy 8074 GACACGTTGGATGAGAGAGAGGACCGGCTCTGGAGTGGCCCGGAGAGACCGGTGCG 8133
Db 7870 CATGACCTCGACAGAGAGAGGCGCGCATCTTGAGCAGGCGCGCGCGCTGCG 7929
Qy 8134 CCAAGCTGGGCGCGAGAGAGAGACTCGGGAGAGGGAGGAAGGCTTGGGGCTG 8193
Db 7930 CCGGCGCTGGCGCGAGAGAGAGCGCTGCGGAGCGGAGGAGGGCGCGCGCTCTG 7989
Qy 8194 GACAGGGGAGAGAGAGAGGCTGCTGAGCAGAGGCGGGTGAAGCTACAGCGCTT 8253
Db 7990 GACGAGGGGAGAGAGAGAGCTGCTGAGCGCGCGGCAAGTGCAGGGCTACAGCGGTA 8049
Qy 8254 TTTGCTGATCTCTGTCAGAGTACCCAGAACTGTACAGAGCGGCCAACAACTTCACTT 8313
Db 8050 CTAAGTACTCTCGTGGAGAGTACCCCGAGTGGCGGAGCGCGCCAAACATCCAGTT 8109
Qy 8314 CATGAGACAGAGGATGGCGCGGAGGTGAC 8345

Db 8110 CTGCGGAGAGAGAGATCGGAGGAGTAAC 8141
RESULT 14
AX250067 LOCUS 8689 bp DNA linear PAT 28-SEP-2001
Sequence 66 from Patent WO0166747.
AX250067 ACCESSION
AX250067 VERSION
AX250067.1 GI:15864503
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 Vernet, C.A., Fernandes, E., Shimkets, R.A., Herrmann, J.L.,
Majumder, K., Macdougall, J., Mishra, V., Mezes, P.S. and Rastelli, L.
Proteins named fctfr and nucleic acids encoding same
Patent: WO 0166747-A 66 13-SEP-2001;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
1..8689
/organism="Rattus norvegicus"
/mol_type="unassigned DNA"
/db_xref="taxon:10116"
ORIGIN
Query Match 40.7%; Score 3397.8; DB 6; Length 8689;
Best Local Similarity 64.8%; Pred. No. 0;
Matches 5469; Conservative 0; Mismatches 2712; Indels 254; Gaps 19;
Qy 35 ATGGACGTGAAGGAGAGAGCTTACCGCTCGCTGCTGACCCCGCGGCGGAGCGGCGG 94
Db 133 ATGGATGTGAAGGATCGCGGACATCGCTTTTGACAGGGACGGTGTGGCAAGAGTGT 192
Qy 95 CCTACACAGCTCGCTCGCGGACAGCGAGGAGGCAAG---CCCCCAGAAAATCGTAC 151
Db 193 CGCTACACAGCTCTCTCTGACAGTGAAGACTGCGGTGTCGCCACGACAGTCTTAC 252
Qy 152 AGCTCAGGAGACCTCTGAAGGCTTACGACGAGGACGCGCGCTAGCCTATGGAGCGGC 211
Db 253 AGTTCCAGTGAGACCTCTGAAGGCTTATGACATGACAGAGATGCACTATGGAACCGA 312
Qy 212 GTCAAGGACATGTGCCGAGAGGCGGAGGAAATTTCTGCCACAGAGTGCACCTTAC 271
Db 313 GTCAAGACCTGGTGACCGGGAGTCCGATGAGTTTCTACAGAGGGCTTAATTAC 372
Qy 272 CTGCGGAGCTGGGCTGGAGAGTAACGCCCTCTACGGGACCTGTACCGGACAGAC 331
Db 373 CTGGCAGAAATGGGAATCTGGAG---CCCTCCCGACACCGAAGTGTACTGTTCCGAC 429
Qy 332 ATTGCGCTGCCCATCGGCTACTCCATGGGGGTGCTCTGATGCCGACATGGAGCT 391
Db 430 ATGGGGATCTCTCCACAGGGCTACTCTCTGAGCCTGGGTCTGATGCGGATCTGGACAC 489
Qy 392 GACAGCGTGTCTCCCTGAGCACCCCGTGGCTCTGTGGGGCGGAGACACAGGTCAGG 451
Db 490 GAGGGAGGATGTCTCCAGAACATGCCATCAGACTGTGGGACGAGGATATAATTCGAG 549
Qy 452 CCGAGCTCTGCTGTCCAGCGCGGCGCAATTCATCTCACACTCACCGACCGAGCAT 511
Db 550 CCGAGCTCTGGGTGTCCAGCGCGGAGAACTCAGCCCTTACTGTGACTGATTTCTGACAA 609
Qy 512 GAAAACACTGAGACTGATCATCCGGCGCGC---TGACAGAACACCGCGCGCTCCGAC 567
Db 610 GAAATAATCGATGACGACAAATGGTGCACCCATTCACCTACATCTCTCTAGCTC 669
Qy 568 GCGCGCGCGCGCTCTCGCAGCGCCCAACCGCCCAACCGACACCGCGGCTCGATTAA 627
Db 670 CTCCATCTGTCTAGCTGCTAGCTTCCCATTAATCTCCACAGTTAGTGTGCC-AGATGCC 728
Qy 628 CTCCCTGAACCGGGCAACTTCAACGCGAGGAGCAACCCCGCGCGG-----674

Db 729 ATTGTAGACACACCTCCATCAGATCATGACACCAACCCCGATGAGAAATTC 788
 QY 675 -----CCCCACGACCACTCGCTCTCCGAGAGCCCCCTCCGCGCGCC 722
 Db 789 CCCTAATTATACCTCTCAGAGATGCTCAGGCGCCCGACGAGCTCCAGTAGTGCC 848
 QY 723 AGGAGCCTGCCACGCCAGAGAACTGGCTCTCAACAGCAACATCCCTCGAGACCA 782
 Db 849 TCCGAAACACACAGACCACTCAACGCTAGAGCCCTCTGCGCACTCTCTATAACCAAC 908
 QY 783 GGAACCTAGGCAAGCAGCACTCTAGGAGCACTGAGCAACACCTCATGAGATGAGCA 842
 Db 909 CCGTCCCAACCACTCTCTGCAACTCTCTCAACAGAACTCACTGACCAATCGCG 968
 QY 843 TTCTCGGCGCTCCCGCA----- 861
 Db 969 GAGTCAAAATCCACGCCCGAGCTCTGCAACCAATGACCTGCGCCACCGCGGAGTCCGT 1028
 QY 862 -----TGATGGGCTTTACAGTGAACGAGGCACTTCCT 891
 Db 1029 TCAGTCCAGGACAGCTGGGTGCTGAACAGTAACTGCGCGCTGAGAGCGGCACTTCCT 1088
 QY 892 CTTCAA-----GCGTAGGACCACTCCCGCTCTTGTGACACCACTCAACAGGATACCACT 948
 Db 1089 CTTCAAGACGCTCTCCGGAAGCACACCCCTCTTCAAGAGCTCTCTCCAGGATACCCCTT 1148
 QY 949 GACGTCCAGCAGTGTACTCTCTCCGCGCGGACCCCTGCGCGAGCACTTCGCGCG 1008
 Db 1149 GACCTCAGGACCGTTTATACACCAACCCCGCTGCTGCCAGGAATACATTCCTAG 1208
 QY 1009 GCGGCGCTTAACTCAAGAACCCCTCAAGTACTGTAACTGGAAGTGCAGAGCCCTGAG 1068
 Db 1209 GAAGGCTTCAAGCTGAAGAACCTCCAAATACTGCAAGTTGGAATGCGCGCGCTGTC 1268
 QY 1069 CGCATCGTATCTCAGCACTCTGCTCATCTCTGTCATCTCTGTCATCTTGTGGCATGCACT 1128
 Db 1269 TGCCATGCGCGCTCTCTCTGCGCAATTTCTGCGCTTATTTATGCAATGCACT 1328
 QY 1129 GTTTGGCCTAACTGGCACTGCAAGCGGATGAGGGGCGAGATGATGAGATCAGGAGGA 1188
 Db 1329 GCTCGGACTCAATTGGCACTCAGCGGCGAGATGGACACACTTTTAA-----ATGG 1382
 QY 1189 CACAGCAGCAAGTGGCTGTGCCAACCGAGCTCTCCCTATAACCCCTCAGGGGCACTGG 1248
 Db 1383 CGTAAGGACCGGCTTACAGGAACGATGATGTGGCAACAGTGCCATCTGGAGGCAAGT 1442
 QY 1249 CTTAGAGACCCCTGACAGGAAGCAAGGAACCAAGGAACCAAGGAAGCCAGTAGTTCTT 1308
 Db 1443 GCGCTGTGTTGAAA----- 1459
 QY 1309 TCCAGAGGACAGTTTTCATAGATTCTGGAGAAATTGATGTGGAAAGGCGAGCCTCCAGAA 1368
 Db 1460 -----ACAGCAGCATAGACAGCGCGAGGAGAGAGTGGTGCAGCGGTGACACAGGA 1511
 QY 1369 GATTCCTCTGCACTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGATCTGAA 1428
 Db 1512 AGTCCCAACAGGGGTGTTTGGAGTCCCAAGATTCAATCAGTCAGCCTCAGTTCTTAAA 1571
 QY 1429 ATTCAATGTGTCTCTGGAAAGGCGCCCTGTTGGCAATTTATGGCAGAAAAGCCCTCC 1488
 Db 1572 GTTCAACATCTCTCTGGGAGAGTGCCTCTTCGGGCTCTACATAAGAGAGAGCTGCC 1631
 QY 1489 TCCTTCAATACACAGTTTGAATTTGGAGTGTCTGAGTGTGGAGGAGGCTCTCTAAACCA 1548
 Db 1632 ACCATCTCATGACAGTATGACTTATGGAACCGCTGGACGAA-----A 1676
 QY 1549 GGAGCGGAGGAGCTAGAGGAGGAGCCCGCGCAGTCTCGGGAGTGTGCCCGCCCTCCAG 1608
 Db 1677 GGAGAGTGGAGTGTGTGAGTCAACAGGAGAGCGGAGAGATCCAGACCTTGGTGA 1736
 QY 1609 CCATGAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTTGGCTTTTACAA 1668

Db 1737 GAACGAGGCTGTGTTCTGTGCACTACTTGGAGTGGGCGCTGTGGCACTCGCTTCTACAA 1796
 QY 1669 TGACGGAAGAGAGTCAAGAGTGGTTTCTTTCTTCTCAACACTGCGATTGAGTCGGTGATAA 1728
 Db 1797 TGACGGAAGAGCAAGAGAGTGTCTCTTCAATACGGTTGTCTTAGATTCACTGTCAGGA 1856
 QY 1729 CTGCCCCAGCACTGATGGCAATGGTGACTGATCTCTGGACCTGCGACTGCTTCT 1788
 Db 1857 CTGTCCACGAACTGCAACGGGAACGGGAAATGGGTGCTGAGCTGTGCTACTGTTTCC 1916
 QY 1789 GGGTTTTCTGGGCGCGACTGTGGCAGAGCCCTCTGCGCCCTGCTCTGTAGGGAAATGG 1848
 Db 1917 AGGATTTCTAGTGTGCACTGGCTTAAAGTGCCTGCTCTTCTGTGAGTGGATGG 1976
 QY 1849 CCAATACATGAAGGAGAGATGCTTGTGCCACAGTGTGCTGGAAGGGCTGTGATGCGATGT 1908
 Db 1977 ACAGTATTTCAAGGAGACATGCCAGTGTACAGTGTGGAAGGAGAGAAATGCCATGT 2036
 QY 1909 GCCACCAACCACTGATGATGTCGCTGAGCAACCATGCACTGCAATCAGGGCAC 1968
 Db 2037 GCCCATGAACCACTGATGATGCTTCTGTGGGGCCAGGCTCTCTGATTTGATGGAA 2096
 QY 1969 CTGCTATCTGCAACCTTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGAATGCAATGACCC 2028
 Db 2097 CTGCGTGTGTGAGCTGGCTACAAGGGCGAGCACTGCGAAGAAAGTGGATTGCTTGCATCC 2156
 QY 2029 CACATGTTTCAAGCGGGGTGTCTGGTGAGAGCGCAATGCCATTGCTTTGTGGGATGG 2088
 Db 2157 AACCTCTCCAGCAATGGTGTCTGTGAAACGAGAGTGTCTATGACGCGCGCTGGGG 2216
 QY 2089 AGGACCAACTCGAGACCCCGAGGCGCACATGCTTTAGACCACTGTTCAGGCGACGGAA 2148
 Db 2217 CGGGCTCAACTCGAGCTGGCGAGGCTCCAGTGCACAGACCACTGTAGTGGCATGGCAC 2276
 QY 2149 CTTCTCTCCGAGACCGGGCTTGTGAGCTGTGACCAAGCTGCACTGGAACAGTGTTC 2208
 Db 2277 TTACTCTCCCTGACTCTGGGCTCTGCAACTGTGATCCGAATTTGGATGGGTCGCACTGCTC 2336
 QY 2209 TATCGAGATCTGTGCTGCGCACTGTGGTGGCCATGGCGTGTGCTAGGGGCGACCTGCG 2268
 Db 2337 TGTGAGTGTGCTCAGTAGACTGTGGCACTCAGCGGCTCTGATCGGGGAGGCTGCGC 2396
 QY 2269 CTGCGAGGATGGCTGATGGGGGAGCTGCGCAACAGGGGCTGCGCACCGCGCTGTGC 2328
 Db 2397 CTGTGAAGAGGCTGAGAGCGCGCTTGTACAGCGCGTGTGCGACCGCTGCTGAT 2456
 QY 2329 CGAGCATGGAGCTCGCGGCGAGCAAGTGCAGTGCAGCCCTGCTGGAATGGCGAACA 2388
 Db 2457 TGAGCACGGGACCTGTAAAGATGGCAATGTGAATGCCGAGAGGCTGGAATGGTGAA 2516
 QY 2389 CTGCAACCATCGCTCACTATCTGGATAGGCTAGTTAAAGAGGGGTTGCCCTGGGTTGTGCAA 2448
 Db 2517 CTGCAACCATTTGAT-----GGCTGCCCTGATTTGTGCAA 2549
 QY 2449 TGGAACGCGCAGATGACCTTAGAAGTGGTTGGGCACTGCTGCTGCGACCTGGGCTG 2508
 Db 2550 CGGTAAACGGGAGATGCACACTGGGTCAGAACAGCTGTGCTGCTGCGAGACCGGCTG 2609
 QY 2509 GAGAGGAGTGGCTGTGACACTTCCATGGAGACTGCTGCGGTGACAGCAAGCAATGA 2568
 Db 2610 GAGAGGCGCGATGCAAGCTTGCATGGAACCTCTGCGGTGATTAACAGGATATGA 2669
 QY 2569 TGGAGATGGCTGTGGAGTGCATGAGCCCTGACTGCTGCTGCCCTCCAGCCCTGTGCAAT 2628
 Db 2670 GGGAGATGGCTGTGGAGTGCCTGAGCCCTGACTGCTGCCCTCCAGTCCAGCTGTGAGAA 2729
 QY 2629 CAACCGCTGCTGCTGCTGCTCCCTTAAACCTCTGACATCATCCAGGAGACACAGTCCC 2688
 Db 2730 GAGCTGCTCTCTGCGGGGTCTCGGAGCCCTTGGAGATCATTCAGCAAGGCCACAGAGA 2789
 QY 2689 TGTGTACAGCAGAACCTACACTCTCTTATGACCGCATCAAGTTCTCTGTTGGGAGGGA 2748
 Db 2790 CTGGCTGCGGTGA-----AGTCTCTTATGATGATCAAGCTCTTGGCAGGCAAGGA 2843

QY 2749 CAGCAGCACATAATCCCGGGAGAACCCCTTTGATGGAGGCATCTTGTGCTATTTCG 2808
 Db 2844 CAGCACCACATCATCTGGAGACACCCCTTCAATAGCAGCCTGGTGTCTGTATCCG 2903
 QY 2809 TGGCAAGTGTAGATCATAGATGAACCCCGCTGGTGTGTGAACATCATGTTTGTCAA 2868
 Db 2904 AGGCCAAGTAGAACCAACGATGGAGCCCGCTGGTGGGTGTGAATGTCTTTTGTCAA 2963
 QY 2869 TAACCCCTCTCTTGGATATACAAATCAGCAGCAAGATGGCAGCTTTGACCTTGTGTACAA 2928
 Db 2964 GTACCCAAATATGCTTACACCATCATCTGCCAGGACGGACAGTTTGACCTGTATGCCAA 3023
 QY 2929 TGGCGGCATCTCCATCATCTCGCGTTTGGAGCGGACACCTTTTATCATCACAGGAGCACAC 2988
 Db 3024 TGGGGCTCTCCCTTGTACTCTTCACTTTGACGAGCCCTTTTATGAGCCGGAGGGCAC 3083
 QY 2989 CTTGTGGCTGCCATGGATCGCTTCTTTGTGTCAGGAACCATCATCATGAGACATGAGGA 3048
 Db 3084 AGTATGGCCGCGGTGAACAGCTTCTATGCCATGGACACCTCGTAAATGAAGACGGAGGA 3143
 QY 3049 GAATGAATTCACAGCTGTGACCTGTGAGCAATTTTGGCCGCCCCCAACCCAGCTCGTCTCTCC 3108
 Db 3144 GAACTCCATCCCGAGCTGTGACCTCAGTGGCTTTGTCCGGCTCATCCGATCATCTCTC 3203
 QY 3109 ATCCCCAGTACGCTCTTGGCAGCTCTGTGCGAGAAAGGCCCATTTGTGCCGGAAAT 3168
 Db 3204 CTCTCCTGTCTCCACCTTCTTACGCGCTTCCCTTGGCGGGAACCCCATTTGTGCTGAGAC 3263
 QY 3169 TCAGGCTTTGAGAGAGAAATCTTATCTCTGGCTGCAAGATGAGGCTGAGCTTACCTGAG 3228
 Db 3264 CCAGGCTTCTATGAGAGATCGAGCTCCCTGGACCAACGATGAGCTTCTTACCTCAG 3323
 QY 3229 CAGCCGACCCCTGGCTACAAATCTGCTGAGATCAGCTCACCCACCCGACCATPCC 3288
 Db 3324 CTCAGAACACAGGCTACAGTCACTGCTGAAGATCACCATGACCTCACCGTCCAGCTGCC 3383
 QY 3289 CTTCAACCTCATGAGTGCACCTCATGCTAGCGGTGGAGGCGCGCTCTTACGGAAGTG 3348
 Db 3384 CTTGAACCTCATCGGGTTCATTTGATGTTGCGGTGGAGGGGATCTCTTCCAGAAATC 3443
 QY 3349 GTTGGCTGCAGCCCGCAGACCTGTCTATTATTTGATTTGGGACAAAGACAGACGCTCTACAA 3408
 Db 3444 GTTCCAGGCTTCTCCAACTGGCTACACATTCATCTGGGACAAAGACAGACGCTTATGG 3503
 QY 3409 CCAGAGGTGTTGGGCTTTCAGAGCCTTTGTTCCGTGGGTATGAAATATGATCCCTG 3468
 Db 3504 CCAAGGGTATTGCGCTATCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3563
 QY 3469 CCCAGATCTAATCTGTGGGAAAAAGAAACACAGTGTCTGAGGGCTATGAAATTCAGCG 3528
 Db 3564 CCCAGTCTCATCTGTGGGAAAAAGAAAGACAGCCCTACTTCAAGGATTCAGCTGAGCCC 3623
 QY 3529 GTCCAGCTTGGAGATGAGCCTTAGACAAACATCATCTGCTTCAACATTCAGAGTGTAT 3588
 Db 3624 TTCCAACTTGGTGGTGGTCCCTGGATAGCAACACACCTCAATGTGAAAGACGGAAAT 3683
 QY 3589 CTTGCACAAAGGAAATGGGAGAACCAAGTTTGTGTCTCAGCAGCTCTCTGTCTATTTGGAG 3648
 Db 3684 ACTACTCAAGGACAGGGGAGAACAGTTCCTGACCCAGCAGCCCGCATCATCACCAG 3743
 QY 3649 CATCATGGCAATGGGCGCGGAGAGCATCTCTGCGCCAGCTGCAACCGCCCTTGTGTA 3708
 Db 3744 CATCATGGTAAACGCTCGCGCAGAGCATCTCTGTCCAGCTGCAATGGCCCTTGTGTA 3803
 QY 3709 CGGCAACAGCTCTGGCCCGCAGTGGCCCTCACCTGTGCTCTGACCGGAGCCTCTATGT 3768
 Db 3804 AGGCAACAAACTGTTGGCCCGCTGGCTGGCTGGGAGATGATGGAGGCTCTTTGT 3863
 QY 3769 GGGTGAATTCACATACATTAAGAGATCTTCCCTCTGCGAAATGTCAACCAACATCTTGA 3828
 Db 3864 CGGTGACTTCAATATATCCGGCGCATCTTCCCTTCTCGAAACGTGACCCAGTATCTTGA 3923

QY 3829 GCTGAGGAATAAGATTTACAGACATAGTCACTAGTCCAGCACAAATACTTACCTGGCCAC 3888
 Db 3924 GTTACGAAATAAGAGTTTAAACATAGCAACAGCCAGGACACAAGTACTACTTGGTGT 3983
 QY 3889 AGACCCCATGATGGGCGCTTCTTCTTCTGACAGCAACAGCCGGGCTTTTAAAT 3948
 Db 3984 GGACCTGTGATGCTGCTCTATGTCTTCTGACCAACAGTCCCGATCTACCGAGT 4043
 QY 3949 CAAAGTCCACTGTGTGTGTGAAGGACCTTTGTCAAGAACTCTGAGTGTGTGGGGAAG 4008
 Db 4044 CAAAGTCTCTAAGCGGAGCAAGACCTTGGCTGGAAATTCGGAAGTGTGGCCGGACTGG 4103
 QY 4009 TGACAGTGCCTCCCTTTGATGACATCTCGTGGGGGATGGTGGAGGCCACAGAGC 4068
 Db 4104 CGAACAAATGTCTACCCCTTTGATGAAGCCCTGTGGGATGGCGGAGGCTGTGATGC 4163
 QY 4069 CACACTCAACCAATCCAGGGGTATTACAGTGGCAAGTTTGGGCTGTACTTCTGTGA 4128
 Db 4164 CACCTGTATGAGCCCTAGAGGTATTGCAGTAGACAGAAACGGGCTTATGATTTTGTGA 4223
 QY 4129 TGGCAACATGATCAGACGCTATCAGATGGATCATCTCCACCTCTGCTGGCTCTAA 4188
 Db 4224 TGGCAACATGATCAGGAGGTCGACCAAAATGGAATCATCTCCACCTGTGGCTCCAA 4283
 QY 4189 TGATCTCATCATCAGCCCGCCACTCAGCTGTGATTCTCTCATGGATATTCCCAAGGTAAG 4248
 Db 4284 TGACCTCAGCTGTCTGACCACTCAGCTGTGATCTTAGCATGAGCGTGGCCAGGTCG 4343
 QY 4249 ACTGAGTGGCCACAGACTTTAGCCATCAACCCAAATGGACAACTCACTTTATGCTCTGA 4308
 Db 4344 TCTAGATGGCCGACAGACCTTGGGTCAACCCATGGACAAATTCCTGTACGCTCTGGA 4403
 QY 4309 CAAAGATGTGCTCTGCAAAATCTCTGAAACACAGCTGGCATTTGCGCGGAGGCC 4368
 Db 4404 GAACAACTGTCTCTGCGGATCAGGAGATCAGGATCAGCATCATCGGGGACGCC 4463
 QY 4369 CATGCACTGCCAGTCTTGGCATTTGACCACTTCTGCTAAGCAAGTGGCCATCCAGC 4428
 Db 4464 CATGCACTGCCAGTCTTGGCATTTGCGCATCTACT --- CGCTCAGCAAGTCTCGCATCTC 4520
 QY 4429 AACCTGGAGTACGACCGCTTGGCTGTTTACAAATGGGCTCTGTATATGCTGA 4488
 Db 4521 TGCTCTGAGTACGACCGCATCTGCAATTTCTCACACCGGGGTGCTCTACATCACCGA 4580
 QY 4489 GACTCATGAGAAAAAGATCAACCGCATCAGGAGTCAACCACTAGTGGAGAGATCTCACT 4548
 Db 4581 GACGACGAGAAAGATCAACCGCTACGCGCTACGCGTCAACCAACGAGAGATCTGCT 4640
 QY 4549 CGTTCCTGGGCGCCAGTGTGACTGTGATAAATGATGCCAACTGTGATTTGTTTTTC 4608
 Db 4641 CTTAGCCGGGCGAGCTCAGACTGTGACTGCAAAAATGACGTCAACTGCATCTGCTATTC 4700
 QY 4609 TGGACCATGTTATGCCAAGGATGCAAGTAAATACCCATCTTCTTGGCTGTGTG 4668
 Db 4701 GGGATGAGCATACGACCGATGCGATTCCTTGAATCTCCCGCTCTCTTGTAGTGTGGC 4760
 QY 4669 TGCTGATGGGAGCTTCTAGCTGGCGACCTTGGGAACTCCAAATCGGTTTATCCGGA 4728
 Db 4761 TCCGATGGCACCCTATCATCGCAGACTCGGGAATATCCGATCAGGGCGGTGAGCAA 4820
 QY 4729 GAAAGCCTTCTTCAACACCCAGAAATGATGAGCTGTCTTCACTCAATTTGACAGGA 4788
 Db 4821 AAACAAACCTGTCTTAAACCGGTTCAACAGTATGAGCTCGCTCTCCGGAGAACAGGA 4880
 QY 4789 GCTCTATCTGTTTGTATACACCGCAAGCACTGTGTACACCCAAAGCTTCCACAGGAGA 4848
 Db 4881 ACTGTACGTTTACCGCGGATGATTCATCAGTACACCGTGGCTGGTGGAGCGGGA 4940
 QY 4849 CTACCTGTACAACTTCACTACACTGGGAGCGGAGATCATCATCATCAGACAGACAA 4908
 Db 4941 GTACTTATACAAATTTACCTACAGCGCTGCAAAATGATGTCCCGAGTTGATGACACAA 5000
 QY 4909 TGGCAACATGGTAAATGTCCGCGGAGACTCTACTGGGATGCCCTCTGGCTGGTGTCTCC 4968

Db 5001 CCGGATCCCTTAAGATCCCGGACAGCAGTGGCATGCCCGACACCTGCTCATGCC 5060
QY 4969 AGATGGCAGGTGTAAGTGGTGAACATGGGACCAACAGTGCATCTCAAGAGTGTGACCA 5028
Db 5061 TGATAATCAGATCATCAACCTTACCGTGGGACCAACCGAGGCTTCAAGCGGTGTCAAC 5120
QY 5029 ACAAGGACACGAGTTGGCCATGATGACATACCATGGCAATTCGGCCCTTCTGGCAACAA 5088
Db 5121 GCAGAACCTGGAGCTGGGCTCATGACTTATGATGGGAACACTGGACTCTTAGCCACAA 5180
QY 5089 AAGCAATGAAACCGATGGACAAACATTTTATGATGACAGCAGCTTTGGCCGCTTGAACAA 5148
Db 5181 GAGCGATGAAACCGGATGGACAACTTTTATGACTATGACACAGGAGGCGCTGTGACAA 5240
QY 5149 TGTGACCTTCCCTACTGCGCAGGTGAGCAGTTTCGAAAGTATACAGACAGTTCAGTGCA 5208
Db 5241 TGTGACTCGCCCCCGGGGTGGTACAGGCTGCGCCGGAAATGGAGAAATCCATCAC 5300
QY 5209 TGTCCAGGTAGAGACTCCAGCAAGGAT- -GATGTACACCAATACCAACCTGTCTGC 5265
Db 5301 CGTTGACATTGAGAACTCCACCGTGATAACGATGTCACTGTGATTACCACTCTTTC 5360
QY 5266 CTCAGGCGCTTCTACACACTGTGCAAGACCAAGTCCGAAACAGCTACTACATCGGGC 5325
Db 5361 AGTGAGGCTCCTACACCGTGTACAAAGTCAAGTGGGAAACAGCTACCACTCTCGAG 5420
QY 5326 CGATGGCTCCTTGGCGCTGTCTGGCCAAAGGATGGAGTGGCGTGCAGACTGAGCC 5385
Db 5421 CAACGGGACCTTGGCGTCACTGACGCCAACGGATGGCGTCACTTCCACAGCGAGCC 5480
QY 5386 COACTGTGCTGGCAGCCGTCAACCCACCGTGGGCAAGAGGAATGTCACTGCGCCAT 5445
Db 5481 CCAAGTCTCGCAGCAACCTCAACCCACCGTGGGCTGTAAACATCTCCCTGCCAT 5540
QY 5446 CGACACGGCTCAACCTGTGGTGGCGCCAGCGCAAGAGCAGGCTCGGGCCAGGT 5505
Db 5541 GGAGAACGCTTGAACCTCACTGAGTGGCGCTGAGGAAGAAACAGATTAAAGCCAAAGT 5600
QY 5506 CACTGTCTTTGGGCGCGCTGCGGGTGCAACCGAAATCTCTATCTTGGACTTTGA 5565
Db 5601 CACCATCTTTGGAGGAGCTTCGGTCCACGGAAGNACTCTGTCTCATGATTATGA 5660
QY 5566 TCGGTAAACGACAGAGAAGATCTATGATGACACCGCAAGTTCACCTTCGGATTCT 5625
Db 5661 CCGAAATATCCGCATGAGAAAGATCTATGACAGCACCGGAAGTTCACCTCGAGATCAT 5720
QY 5626 GTACACAGCGGGCGGCCAGCCTCTGTGTACCCAGCAGCAGGCTGAATGTGTCAA 5685
Db 5721 TTATGACAGGTGGCGCCCTTCTGTGTGCTCCAGCAGTGGACTGGCGGCCGTCAA 5780
QY 5686 CGTGACATPACTCCCTGGGGTTACATTTGCTGGCATCCAGGGGGCATCATGTGTGAAG 5745
Db 5781 TGTCTCTTCTTCAACGGCGCCTGGCGGCTCCAGCGGGGCCATGAGCGAGAG 5840
QY 5746 AATGGAATACACAGCGGCGCGCATCATCCAGGATCTTGCCTGATGGGAAGACATG 5805
Db 5841 GACAGACATTGACAAAGCCCGGATGTGTCCCGAAATGTTCGCCAGCGGGAAGTCTG 5900
QY 5806 GAGCTACACATPACTTAGAGAAGTCCATGGTGTCTACTACTACAGCCAGAGGAGATAT 5965
Db 5901 GAGCTATTCTTACCTTGACAACTCCATGGTCTCTCTCTGTCAGAGCCAGCGTCACTAT 5960
QY 5866 CTTTGAGTTGCAAGAAATGACCGCTCTCTTTGTGACGATGCCAACCGTGGCGGCA 5925
Db 5961 ATTTGAATATGACTCTCTGTGACCGCTCCAGCGAGTCAACATGCCCCAGTGTGCGCGCA 6020
QY 5926 GACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCTGAGG 5985
Db 6021 CAGCATGTCACGACACCTCCATGGCTATCATCCGGAACATTTAAACCCACCGGAAG 6080
QY 5986 CAATGCTCTAGTATACAGGACTTCACTGAGGATGGGACCTCTTCCACACCTTCTACCT 6045

Db 6081 CAACGGCTCGGTCACTTTGACTATAGTAGCGGCCGATCCTGAAGACGCTCTTTCTCT 6140
QY 6046 GGGCACTGGCCGAGGGTGATATACAAAGTATGCAAACTGTCAAAGCTGCACAGACGCT 6105
Db 6141 GGGCACCGGCGCCAGGTGTCTTATAAGTACGGAATACTGTCCAAAGTATTCGAGATCGT 6200
QY 6106 CTATGACACCAACAAAGTTCAGTTTCACTATGACGAGACGGCAGGATGCTGAGACCAT 6165
Db 6201 CTACGACGACACTGCGCTCACCTTCGCTATGACGAGACCACTGGCGTCTCTGAAGATGGT 6260
QY 6166 CAACTTACAGAAATGAGGGCTTCACTTGCACCACTCCGCTACCGTCAGATTGGGGCCCTGAT 6225
Db 6261 GAATCTCCAAAGCGGGGCTTCTCTGTGTACCATCAGTACCGAAAGTTCGGGCCCTCGT 6320
QY 6226 TGACCGACAGATCTTCGGCTTCACTGAGGAAGCATGTCAACGCCGCTTTTGTACTACA- 6284
Db 6321 GGACAAAGCATTTACAGGTTCTCTGAGGAAGCATGATCAACGCCAGGTTTGGATTACAC 6380
QY 6285 - -ACTATGACAAACAGCTTCGGGTGACCAAGTCAAGGCTGTGATCAACAGACCCCACT 6342
Db 6381 CTACGACGACAGCTTCGGCATCGCAGCATCAAGCCGCTCATCAGTGAGACTCCCT 6440
QY 6343 GCCATTGATCTCTATGCTATGATGTGTGAGCAAGACAGACAGCTTTGGGAAGTT 6402
Db 6441 TCCCGTTGACCTCTACCGCTACGATGAGATTTCTGCAAGGTGGAACACTTCGGCAAGTT 6500
QY 6403 TGGTGTCAATTACTATGACATTAACAGATCATCAACAGCTGTCTATCAACCAACAA 6462
Db 6501 CGGGGTCACTACTACGACATCAACAGATCATCACTCACTGCGCTCATGACACTCAGCAA 6560
QY 6463 GCATTTTGTATGATATGGCAGGATGAAGAAAGTGAAGTATGAGATCTCCGCTCGCTCAT 6522
Db 6561 GCATTTTGCACCCATGGGCGCATCAAGGAAGTGCAGTATGAGATGTTCCGGTCCCTCAT 6620
QY 6523 GTACTGATGACGCTCCAGTATGATTAACATGGGGGAGTAGTGAAGAGAGCTGAAGT 6582
Db 6621 GTACTGATGACGGTGCAATATGACAGTATGGGCGGGTTCATCAAGAGGAACTGAACT 6680
QY 6583 AGGACCTACGCAATACCACTCGCTACTCTATGATGATGATGCTAGCGGCGAGCTGCA 6642
Db 6681 GGGGCCCTATGCCACACCAACAAAGTACACTATGACTACGAGGGGACGGCAGCTCCA 6740
QY 6643 GACAGTCTCCATCAATGACAAAGCCACTGTGGCGCTACAGTACGACCTCAATGGGAACCT 6702
Db 6741 GAGTGTGGCGCTCAATGACCGGCTACCTGGCGTTATAGCTATGACCTCAATGGGAACCT 6800
QY 6703 GCACCTACTGAGCCCTGGGAACAGTGCAGGCTCAACCACTACGCTATGACATCGGCA 6762
Db 6801 GCACCTGTAAACCCAGGAACAGTGTCTGCTCATGCGCTTACGCTATGACCTCCGTA 6860
QY 6763 CCGCATCACTCGGCTGGGTGACGTGCAATACAAAGATGGATGAGGATGGCTTCTGAGCA 6822
Db 6861 CCGGATAACCAAGCTAGGGGACGTGCAGTACAAATCGATGATGATGGCTATTTATGCCA 6920
QY 6823 GCGGGCGGTGATCTTTGAGTACAACTCAGCTGGCTGCTCATCAGGCCCTACACCG 6882
Db 6921 GAGAGGATGTGATCTTTGAATACAACTCCAAGGCGCTTCTAACGAGAGCGTACAAAC 6980
QY 6883 GGTGCGAGCTGAGGTGTGAGTACCGCTACGATGCGCTGGGGCGCGGTGTCCAGCAA 6942
Db 6981 GGGCAGCGGTGAGTGTGAGTACCGCTATGATGGCGTGAGCCCGGGCTTCTTACAA 7040
QY 6943 GAGCAGCCACAGCCACCACTGAGTTCTTATGACAGCTGACCAACCCCAAGGT 7002
Db 7041 GACCAACCTGGGCCACCACTACAGTACTTCTATTCCGACCTCCACCAACCCACACGAT 7100
QY 7003 CACCACCTGTAAACCACTCCAGCTGAGATCACTCCCTCTACTACTAGCTTCAAGG 7062
Db 7101 CACCATGTTTAAACCACTCCAACTTGAGATCACTCACTTACTATGACTCCAGGG 7160
QY 7063 ACACCTCTTTGCGCTGAGCTGAGCAGTGGTGTGATGAGTTTACATAGCTTTGACAAAT 7122
Db 7161 CCACCTCTTTGCGCTGAGAGCAGTAGTGGGGAAGAGTACTATGTTGCTCAGATAACAC 7220

QY 7123 CGGACCCCTCTTGTGCTTTTGTAGTGAACAGGTTTGTATGATCAAGCAAAATCTGTACAC 7182
 Db 7221 CGGACTCTCTGCTGCTTTTGTATATCAATGGCTCATGATCAAGCAATCTCAATACAC 7280
 QY 7183 AGCTATGGGAGATCTACATGATACCAACCCCAATTTTCAGATCATATAGGCTACCA 7242
 Db 7281 AGCTATGGGAGATTTACTATGATCTCAATCCAGACTTTTCAGATGCTATCGGCTTCCA 7340
 QY 7243 TGGTGGCTCTATGATCCACTCACCAAGCTTGTCCATCGGCGGAGATTATGATGT 7302
 Db 7341 CGGAGGCTCTACGACCCCTTACCAAGCTCGTTCACTTTACGAGCGTATATGACGT 7400
 QY 7303 GCTGCCGAGCGCTGGATAGTCCAGACCAAGAGCTGTGAAGCACCTTATAGACAGAA 7362
 Db 7401 GCTGGCAGGCGGTGGAGCTCCCGGACTACACCATGTGGAGAA---TGTGGGCAAGGA 7457
 QY 7363 CGTCATGCTTTTAACTCTATATGTTCAAAACCAACCCCATCAGCAACTCCAGGA 7422
 Db 7458 GCCAGCCCTTCAACCTGTATGTTCAAGAACCAATCCACTAGTAAAGCTGGA 7517
 QY 7423 CATCAAGTCTTCTATGACAGATGTTAAAGCTGTCTCACCTTTGGATTCCAGCTACA 7482
 Db 7517 TTTAAAGACTACGTGACAGAGCTGAAGAGCTGCTGTGATGTTTGATTTCAGCTCAG 7577
 QY 7483 CAAGTATCCCTGTTATCCCAACCAAGCATGGATGCCATGGAACCTCTACAGCT 7542
 Db 7578 CAACATCTCTGCTGATTTCCCAAGAGCCAAATGATTTTGTGCTCCCTCCCTATGA--- 7634
 QY 7543 CATCCACACAGATGAAGCGAGGAGTGGGCAACAGCAAGTCTATCTCCGGGTACA 7602
 Db 7635 -----ACTGTCAGAGACCAAGCAAGTGAAGATGAGACAGCTCAATTACAGCTTCCA 7685
 QY 7603 GTGTGAAGTACAGAGAGCTCAAGGCTTTGTCACTTTAGAACGTTTGAACAGCTCTA 7662
 Db 7686 GCAGACAAAGAGAGGCAATAACAGGCTTTTCTGGCTCTAGAGGACAGCTCATCTTAA 7745
 QY 7663 TGGCTCCACATCACCAGCTCCAGAGGCTCCAAAGCAAGATTTGCATCCAGCG 7722
 Db 7746 AAGCTCCATGAGGATCCAGAGAAAGC-----AGGCCACTGTTGCTAGACCAC 7799
 QY 7723 CTCAGTCTTTGGCAAGGGGCTCAAGTTTTCCTTGAAGAGTGGCCAGTGAACACAGACAT 7782
 Db 7800 GCCCATCATCGCAAGGACATCATGTTCCGCATCAAGAGAGGCGGTGACCAAGCGGT 7859
 QY 7783 CATGATGTGCCAATGAGATGGGAGGGTTGCTGCCATCTTGAAACCATGCCACTA 7842
 Db 7860 GTCTAGCATCGCAGTGAAGGACGCGCAAGGTAGCATCCGTGTGAAACACGCTTACTA 7919
 QY 7843 CCTAGAGAACCTGCATTTACCATTTGATGGGTGGATACCCATTTGTTGAACACAGG 7902
 Db 7920 CTTGACAGATGCTACTACAGATCGAGGCAAGGACACACACTACTTCTGTAAGATCG 7979
 QY 7903 ACCTTCAGAGATGACTGGCCATCTCTGGGCTCTAGTGGGGGGCGGCAACCTTGAAGAA 7962
 Db 7980 TGCAGCGACGCTGACCTCTGTTTACGCTGGGACCAACCATTTGGGGCGAAGTGTGAGAG 8039
 QY 7963 TGGGCTCAACGTCATGTTGCCAGATCAACACAGTACTTAAATGGCAGGACTAGACCTA 8022
 Db 8040 CGGGGTGAACGTGACCGTGTACACCCACGCTGCTGTGTGAACGCGAGGACTCGAAGTT 8099
 QY 8023 CACAGACATCCAGTCTCCAGTACGGGGCACTGTGTTGAACACAGCTACCG----- 8073
 Db 8100 CACCAACATTGAATTCAGTACTCCACGCTGCTGTCTAGCATAGCTACGCGCTTCACCCC 8159
 QY 8074 GACACGTTGATGAGAGAGGACGCGGCTCTCTGAGCTGGCCCGGAGAGCGGCTGG 8133
 Db 8160 GCACACATGATGATGAAGAGAGGCGCGGCTCTCTGACCAACGCGGACAGAGGGCTGG 8219
 QY 8134 CCAAGCGTGGGCGCGGAGCAGCAGACTGCGGGAAGGGAGGAGGAGGAGGAGGAGGAGG 8193
 Db 8220 TACTGCTGGGCGCAAGGAGGACAGCAAGAGCCAGGAGCGGAGAGGAGGAGGAGGAGGAG 8279

QY 8194 GACAGAGGGGAGAGAGCAGCAGGCTGCTGAGCAGAGCGCGGTGCAAGGCTACGAGCGCTT 8253
 Db 8280 GACGAGGGCGGAGAGCAGCAACTCTTGAGCAGCGGAGCGGTGCAAGGCTTATGAGGCTA 8339
 QY 8254 TTTCTGTGATCTCTGTCGAGCAGTACCCAGAACTGTTCAGACAGCGCCCAACACATCCACTT 8313
 Db 8340 TTACGTGCTTCCGTTGGGAACAGTACCCAGAGCTGGCAGACAGTACGAGCAATCCAGTT 8399
 QY 8314 CATGACAGAGCAGAGATGCGGCCGAGGTGACAGA 8348
 Db 8400 CTTAAGACAGATGAGATGCGAAGAGGTAACAAA 8434
 RESULT 15
 AF086607
 LOCUS
 DEFINITION
 Rattus norvegicus neurestin alpha mRNA, complete cds.
 ACCESSION
 AF086607
 VERSION
 AF086607.1 GI:5712200
 KEYWORDS
 Rattus norvegicus (Norway rat)
 SOURCE
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 REFERENCE
 1 (bases 1 to 8689)
 Otaki, J.M. and Firestein, S.
 Neurestin: putative transmembrane molecule implicated in neuronal
 development
 Dev. Biol. 212 (1), 165-181 (1999)
 JOURNAL
 MEDLINE
 99350226
 PUBMED
 10419693
 REFERENCE
 2 (bases 1 to 8689)
 Otaki, J.M.
 Direct Submission
 Submitted (25-AUG-1998) Department of Biological Sciences, Columbia
 University, Fairchild Center, New York, NY 10027, USA
 JOURNAL
 FEATURES
 1. 8689
 Location/Qualifiers
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /strain="Sprague Dawley"
 /db_xref="taxon:10116"
 /tissue_type="olfactory bulb"
 133..8430
 /note="putative neuregulin-like receptor; proposed as a
 receptor for odorant receptors; may function for target
 recognition and synapse formation; similar to Homo Sapiens
 gamma heregulin, Drosophila Ten(a), mouse DOC4, and
 gamma, and delta identified"
 /codon_start=1
 /product="neurestin alpha"
 /protein_id="AAD47383.1"
 /db_xref="GI:5712200"
 /translation="MDVKDRHRSILTRCCKEYRTSSSLSDSDCRVPTOKSVSSSE
 TLKAYDHSRHYGNRVDLVRHREDFSRQANFTLAEGLCEPSPHRSYCDMGI
 LHQYSLSTGSDADSTEGMSPEHAIRLWGRGIKSRSSGLSRENSALTDSNE
 NKSDNCRPTPTSSSLLPSAQLPSHPNPVSCQPLDLSNTSHQIMNDPEEF
 SPNSVLLACSGPQOASSSGPNHHSQTLRPLPPPHNHTLSHHSSANSLSLT
 NRSQTHAPAPNDLATPSSVOLQDSVLNSNVPLETRHFLFKTSSGSLPSSSS
 PGYITGLVTTPRLLPRNTFSKAPKLKPKSKYCKWCAALSAIAAAILLALLA
 YETVHLHLGLQLPADGHTFNNGRTGLFGNDVATVPGGKVPWSLKNSSDSCG
 AEVGRVTVQEPVPPVFWRSQIHISQPKFNISLGKDALFVIRGLPSPHQIDP
 MERLDKPKWSVESPERRSIQTLVQNEAVFVQYLDVLMELAFVIRGLDPEKWFN
 TVLSDVODCPNCHGECVSGCHCPPLGLADCAKACPVLCGNGQYSKGTCCQ
 YSGWAECDVPMNOCIDPCGCHGSCIDGNCVCAAGYKGBEHCDEVDLDTCCSHGV
 CVNGCLGSPKGLNCLARVQCPDQCSGHCTYLPDSGLCNCDPNMWPSPCSEVCS
 VDCGTHGVCISACRCBEGWGAACDQVCHRCIEHCTCKDKECEGEGNGHCTI
 DCCPLCNGNGRGTGQNSWCQCTQWRGFGCQNVAMETSCADKNKDEGDLVCLDP
 DCLLSAQCNLLCGRSDPLDIQQGTQDPAVKSFYDRIKLLAGKGSALTLLH
 FNSLSVLRIGQVTTDGTPLGVNVSFVKYGYTTITRODPTFLIANGGSALTLLH
 PERAPFMSRRTVMPMNSFYAMDITLVMKTEENSIPSCDLSGFRVDPFIILISPLSTF

Db 1677 GGAGAAAGTGGAGTGTGGTGGAGTCAACCCAGGAAACGCCGAGCATCCAGACCCCTGGTGCA 1736
Qy 1609 CCATGAGACAGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCAGCTTGGCTTTTACAA 1668
Db 1737 GAACGAGGCTGTGTTGTGCGAGTACTTTGGATGTGGGCTGTGGCAGCTCGCCCTTTACAA 1796
Qy 1669 TGACGAAAGAGTCAAGAGTGGTTCCTTTCTTCCACCACTGCCATTGAGTCGGTGATAA 1728
Db 1797 TGAGGCAAGGACAAGGAGATGGTCTCTCTCAATACGGTTGTCTTAGATTCAAGTGCAAG 1856
Qy 1729 CTGCCCCAGCAACTGCTATGCAATGTGAGTGTGATCTCTGGGACCTGCCACTGCTTCCT 1788
Db 1857 CTGTCCACGAACTGCCACGGGAACGGGAATGGTGTCTGGACTGTGTCACTGTTTCCC 1916
Qy 1789 GGGTTTCCTGGGCCCGGACTGTGGCAGAGCCCTCTGCCCGTGTCTGTAGCGGAATGG 1848
Db 1917 AGGATTCCTAGTGCAGACTGCGCTAAAGCTGCTGCCCTGTCTGTGCGATGGGAATGG 1976
Qy 1849 CCAATACATGAAGCAGATGCTTTGTGCCACAGTGGCTGGAAAGCGCTGAGTGGGATGT 1908
Db 1977 ACAGTATTTCCAAAGGACATGCCAGTGTCTACAGTGGCTGGAAAGGAGCAGAAATGCCATGT 2036
Qy 1909 GCCACCAACAGTGTATGATGTGGCTGTGGCAGCAACCATGGCACCTGCATCAAGGSCAC 1968
Db 2037 GCCCATGAACAGTGTCAATCGATCTCTTCTGTGGGGGCCACGGCTCTCGAATGATGGAA 2096
Qy 1969 CTGCATCTGCAACCCCTGGCTACAGGGCGAGAGTGTGAGGAAGTGAAGTGCATGGACCC 2028
Db 2097 CTGCGTGTGTGAGCTGGCTACAGGGCGAGCACTGGGAAGTGGATGCTTGGATCC 2156
Qy 2029 CACATGTTACGGCGGGGTGTCTGGTGTGAGAGGCGAATGCCATTGCTTTGGGATGGG 2088
Db 2157 AACCTGTCTCCAGCCATGTGTGTGTGAACGGAGAGTGTCTATGACGCCCGGCTGGG 2216
Qy 2089 AGGCACCAACTGCCAGACCCCGAGGGCCACATGCTTAGACCAAGTGTTCAGGCCACCGAAC 2148
Db 2217 CGGGCTCAACTGCGAGTGGCGAGGGTCCAGTGGCCAGACCAAGTGTAGTGGCATGGCAC 2276
Qy 2149 TTCTCTCCGAGACACCGGGCTTTGAGCTGTGACCCAGAGTGGACTGGACAGACTGTTCC 2208
Db 2277 TTACCTCCCTGACTCTGGCCCTCTGCAACTGTGTATCCGAATGGATGGGTCCCGAGTGTCTC 2336
Qy 2209 TATCGAGATCTGTGTCGCCAGTGTGGTGGCCATGGCGTGTGCTAGGGGAGCACTGCCG 2268
Db 2337 TGTGAGTGTGCTCAGTAGACTGTGAGCTCAAGGGGTCTGCAATCGGGGAGCTGCGG 2396
Qy 2269 CTGCGAGGATGGCTGGATGGGGGAGCTGTGGACAGCGGGCTGTGCCACCGCGCTGTGC 2328
Db 2397 CTGTGAAGAGGGCTGGACAGCGGGGTGTGTGACAGCGGCTGTGCCACCCCGCTGCAT 2456
Qy 2329 CGAGCATGGGACCTGCCCGCAGGCAAGTGGAGTGGAGCCCTGGCTGGGAATGCGGAACA 2388
Db 2457 TGAGCAGCGGACCTGTAAAGATGGCAATGTGAATGCGAGAGGCTGGAAATGTGAACA 2516
Qy 2389 CTGCACCATCGCTCACTATCTGGATAGGGTATGTTAAAGAGGGTTGCCCTGGGTTGTGCAA 2448
Db 2517 CTGCACCATTGAT-----GGCTGCCCTGATTGTGCAA 2549
Qy 2449 TGGCAACCGCAGATGCTTAGACCTGAATGGTTGGCATGGCTGTGCCAGCTGGGCTG 2508
Db 2550 CGGTAAACGGGAGATGCACACTGGGTCAAGACAGCTGGCAGTGTCTGTGCCAGACCGGCTG 2609
Qy 2509 GAGAGGAGCTGGCTGTGACACTTCATGGAGACTGCTGTGGGTGACACCAAGACATGA 2568
Db 2610 GAGAGGCGCCGAGTGAACAGTGTGCCATGGAAACCTCTGCGCTGATAACAGGATATGA 2669
Qy 2569 TGGAGATGGCTGTGGTGAATGCAATGGACCTGACTGTGCTCCAGCCCTGTGCCATAT 2628
Db 2670 GGGAGATGGCTGTGGTGAATGCAATGGACCTGACTGTGCTCCAGTCAAGTCAAGTGTGAGAA 2729
Qy 2629 CAAACCGCTGTGCTGTGCTCCCTTAAACCTCTGGACATCATCCAGGAGACACAGGTCCC 2688
Db 2730 CAGCTGTCTGTGCGGGGTCTCGGGACCCCTTGGACATCATTCAGCAAGGCCAGACAGA 2789

Qy 2689 TGTGTCAAGCAGAAACCTACACTCTCTTATGACCGCATCAAGTTCTCTCGTGGCAGGGA 2748
Db 2790 CTGGCTCGGGTGA-----AGTCTCTTATGATCGTATCAAGCTCTTTGGCAGGCAAGGA 2843
Qy 2749 CAGCAGCAGCATATATCCCGGGGAGAACCCCTTTGATGGAGGATGCTTGTGTTATTCG 2808
Db 2844 CAGCAGCAGCATATATCTCTGGAGACACCCCTTCAATAGCAGCCTGTGTCTCTGATCCG 2903
Qy 2809 TGGCAAGTGTATGACATCAGATGGAAACCCCTGTGTGTGTGAACATCACTGTTTGTCAA 2868
Db 2904 AGGCCAAGTAGTAACACAGGATGGACCCCTGTGTGTGTGTGAATGTCTTTTGTCAA 2963
Qy 2869 TAACCTCTCTTTTGGATATACAATCAGCAGCAAGATGGCAGCTTTTGAATTTGGTGACAAA 2928
Db 2964 GTACCCAAATATGCTTACACCATCACTCGCAGGACGGCAGCTTTGACCTTGATTGGCAA 3023
Qy 2929 TGGCGGCATCTCCATCATCTCTGGGTTCGAGGGGACCTTTTTCATCACAGGAGCAGCAC 2988
Db 3024 TGGGGGCTCTGCTTGTGATCTCTTCACTTTGAGGAGCCCTTTTATGAGCCGGAGCGCAC 3083
Qy 2989 CCTGTGGTGGCATGGGATCGCTTCTTTGTCTATGGAACCAATCATCATGAGACATGAGGA 3048
Db 3084 AGTATGGCGCGGTGGAACAGCTTCTATGCCATGGACACCCCTGGTAAATGAAGACGGAGGA 3143
Qy 3049 GAATGAGATTCACAGCTGTGACCTGAGCAATTTTGGCCCGCCCAACCCAGTCTGTCTCC 3108
Db 3144 GAATCCATCCCAAGCTGTGACCTCAGTGGCTTTGTCCGGCTTATCCGATCATCATCTC 3203
Qy 3109 ATCCCACTGAGCTCTTCCGCGAGCTCTGTGCAAGAAAGGCCCCCATTTGTGCCGGAAT 3168
Db 3204 CTCTCTCTGTGCACTTCTTTCAGGGCTTCCCTTGGCGGAACCCCATTTGTGCTGAGAC 3263
Qy 3169 TCAGCTTTGAGGAGGAATCTCTATCTGTGGTGTGAAGATGAGGCTGAGCTGAGTGTAG 3228
Db 3264 CCAGGTTTCTATGAGGAGATCGAGCTCCCTGGCACCACAACTGAAGCTCCGTTACCTCAG 3323
Qy 3229 CAGCCGAGCCCTGGCTACAAATCTGCTGAGGATCAGCTCACCCACCCGACCATCCC 3288
Db 3324 CTCCAGAACAGCAGGATACAGTCTACTGTGAAGATCACATGACCCAGTCCAGGTTGCC 3383
Qy 3289 CTTCAACCTCATGAAGTGTCACTCATGCTAGCGGTGGAGGGCCGCTTCTTCAAGAAATG 3348
Db 3384 CTTGAACCTCATTCGGGTTCACTTGTGTTCCGCTGGAGGGCATCTCTTCCAGAAATC 3443
Qy 3349 GTTCGCTCAGCCCGCAGACCTGTCTATTTTCAATTTGGGACAGACAGACGCTCAAA 3408
Db 3444 GTTCAGGCTTCTCCAACTGGCCCTACACATTCATCTGGGACAGACAGACGCTTATGG 3503
Qy 3409 CCAGAAAGTGTTTGGGCTTTCAGAAAGCTTTTTCGTTGGGTTATGAATATGAATCTG 3468
Db 3504 CCAAGGGTTTATGSCCTATCGGATGCTGTGTCTGTGATTTGAATATGAGACCTG 3563
Qy 3469 CCCAGATCTATCTGTGGGAAAAAGAAACAGTGTCTGAGGGCTATGAATATGACGC 3528
Db 3564 CCCAGTCTCATCTCTGTGGGAAAAAGGACAGCCCTACTTCAAGGATTCGAGCTGGACCC 3623
Qy 3529 GTCCAAAGTGTGGAGATGGAGCTTAGACAAACATCATGCCCTCAACATTTCAAAGTGTAT 3588
Db 3624 TTCCAACTTGTGTGCTGTGCTTCCCTGGATAGCAACACACCTCAATGTGAAAGCGGAT 3683
Qy 3589 CTTGCAAAAGGAATGGGAGAACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3648
Db 3684 ACTACTCAAGGACAGGGGAGAACAGTGTCTGACCCAGCAGCGCCGCTCATCACACAG 3743
Qy 3649 CATATGGCATGTGGCCCGGAGAGCATCTCTGTGCCCGCAGCTGCAACGCTTGTGTGA 3708
Db 3744 CATATGGGTAAAGTGTCCCGGAGAGCATCTTCTGTCTCCAGCTGCAATGAGCTTGTGTGA 3803
Qy 3709 CGGCAACAGCTCTGGCCCGCAGTGGCCCTCACTGTGTGGCTTGTGACGGGAGCCTTCTGTGT 3768
Db 3804 AGGCAACAAACTGTGTGGCCCGCTGGCTGTGGGGATCGATGGGAGCCTCTTGTGT 3863

3769 GGGTGAATTTCACTACATTAGAGGATCTTCCCTCTCGAAATGTACCAACATCCTAGA 3828
3864 CGGTGACCTCAATTATATCCGGCGCATCTCCCTCTCGAAACGTGACCAATCTTGA 3923
3829 GCTGAGGAATAAGATTTACAGACATAGTACAGTCCAGCACACAAAATACCTGCGCAC 3888
3924 GTTACGAATAAAGAGTTTAAACATAGCAACAGCCAGGACACAAAGTACTACTTGGTGT 3983
3889 AGACCCCATGAGTGGGGCGCTCTTCTTCTGACAGCAACAGCCGGGGGCTTTAAAT 3948
3984 GGACCCCTGTGACTGGCTCGCTATGTCTCTGACACCAACAGTGC CGGATCTACCGAGT 4043
3949 CAACTCCACTGTGGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTGGGGGACAGG 4008
4044 CAAAGTCTCTAAGCGGAGCAAGACCTGGCTGGGAAATTCGGAAGTTGTGGCGGAGCTGG 4103
4009 TGACCAAGTCCCTTCTTGATGACATCGCTGGGGGATGCTGGGAAGGCCACAGAGC 4068
4104 CGAACATGTCTACCCCTTGTAGAACCCGCTGTGGGATGGCGGAAGGCTGTGGATGC 4163
4069 CACACTCAACAAATCCAGGGGTATTACAGTGGCAAGTTTGGGTGATCTACTTCTGTTGA 4128
4164 CACCCCTGATGAGCCCTAGAGGTATTGCAGTAGACAAAGACGGGCTTATGTTATTTGTTGA 4223
4129 TGGCACCATGATCAGACCATCGATCAGAAATGGGATCATCTCCACCTGCTCGCTCTAA 4188
4224 TGCCACCATGATCCGGAAGGTGACCAAAATGGAAATCATCTCCACCTGCTGGGCTCAA 4283
4189 TGATCTCACATCAGCCCGGCCACTCAGCTGTGATTCTGTCTATGATATTTCCCAAGTAA 4248
4284 TGACCTCAGAGTGTCCGACCACTGAGCTGTGACTCTAGCATGAGCTGGCCGAGTCCG 4343
4249 ACTGAGTGGCCACAGACTTAGCCATCAACCAATGGACAACTCATTTATGCTCTGA 4308
4344 TCTAGAAATGGCGCAGACACTTGGGTCAACCCCATGGACAATTTCCCTGTACGTCTGGA 4403
4309 CAACAATGTGTCTGCAAAATCTCTGAAACACACAGGTGGCATTTGCGCGGAGGCC 4368
4404 GNAACAAGTCTATCTGCGATACCGAGAAATCAACAGGTGAGATCATCGCGGAGCGCC 4463
4369 CATGACCTGCCAGGTCTCTGGATTTGACCACTTCTGTAAGCAAGTGGCCATCCACGC 4428
4464 CATGACCTGCCAGGTCTCTGGATTTGACCACTTCTGTAAGCAAGTGGCCATCCACGC 4488
4429 AACCTGTGAGTCAGCCCGCTTTGGCTGTTTACACAAATGGGTCTGTATATGTTGA 4500
4521 TGCTCTGAGTCAGCCGCGCATCGCCATTTCTACACCGGGGTGCTCTACATCACCGA 4580
4489 GACTGATGAGAAAGATCAACCGCATCAGGAGTCAACCACTAGTGGAGAGATCTCACT 4548
4581 GACGACGAGAAGAGATCAACCGCTTACCGCATGAGTCAACCACTAGGAGATCTGCT 4640
4549 GGTGCTGGGCGCCAGTGGCTGTGACTGTAAATGATGCCAACTGTGATTTGTTTC 4608
4641 CTTAGCCGGGCGAGCTCAGACTGTGACTGCAAAATGACGTCAACTGCATCTGCTATTC 4700
4609 TGGAGACGATGTTATGCAAGGATGCAAGTAAATACCCCATCTTCTTGGGTGTTG 4668
4701 GGGAGATGACGATACGCAAGATGCAATCTTGAATCCCGCTCTCTTACTGCTGGC 4760
4669 TGCTGATGGGAGCTCTACGTGGCGACCTTGGGAACATCGAAATCGGTTTATCCGGA 4728
4761 TCCGATGGCACCATCTACATCGCAGACCTCGGAATATCCGATCAGGCGGTGAGCA 4820
4729 GAACAAGCTTTCTCRAACCCAGAACATGATGAGTGTCTTCAACCAATGACAGGA 4788
4821 AAACAAACCTGTTCTTAAAGCGTTCAACAGTATGAGGCTCGCTCTCCGGAGAACAGGA 4880
4789 GCTCTATCTGTTGATACCCGCAAGCAGCTGTACACCCAAAGCCTGCCCAAGGAGA 4848
4881 ACTGTACGTGTTCAACCGGATGTTATCCATCAGTACACCGTGGCTGGTGAACCGGGA 4940
4849 CTACCTGTACAACTTCACTGAGGAGCGGCAATCACTCATCACAGAACAA 4908

4941 GTACTTATACAAATTTCACTACAGCGCTGACAATGATGTCAACGAGTTGATTGACAA 5000
4909 TGCAACATGTTAAATGTCCGGCAGACTCTACTGGGATGCCCTCTGGCTGGTGTCC 4968
5001 CGGAATTCCTTAAAGATCCGGCGGACAGCATGGCATGCCCGACACCTGCTCATGCC 5060
4969 AGATGGCCAGGTGTACTGGGTGACCATGGCCACCAACAGTGCACCTCAAGAGTGTGACCC 5028
5061 TGATAATCAGATCATCACCTTACCGTGGGACCAACAGGAGGCTCAAGACCGGTGTCAAC 5120
5029 ACAAGACAGAGTGTGGCCATGATGACATACATGCAATTCGGGCTTCTTGGCAACCAA 5088
5121 GCAGAACTGGAGCTGGGCTCTATGACTTATGATGGAAACACTGGGACTCTCTGACCCAA 5180
5089 AAGCAATGAAACCGGATGGACAACTTTTATGATGACGACAGCTTTTGGCGGCTGACAA 5148
5181 GACGATGAAACCGGATGGACAACTTTTATGACTATGACAGGAGGCGCTGACCAA 5240
5149 TGTGACCTTCCCTPACTGGCCAGGTGAGCAGTTTCCGAAGTATACAGACGTTCAGTGCA 5208
5241 TGTGACTCGCCCAACCGGGGTGTGACAGCCTGACCGGAAATGGAGAAATCCATCAC 5300
5209 TGTCCAGGTAGACCTCCAGCAAGAT--GATGTACCAATAACCAACCTGCTCTCC 5265
5301 CGTTGACATTGAACTCCAAACCGTGAACGATGATCACTGTGATTACCAACCTCTCTTC 5360
5266 CTGAGCGGCTTCTACACACTGTGCAAGACCAAGTCCGGAACAGCTACTATCATCGGGC 5325
5361 AGTGAGGCTCTCTACACCTGTGTACAAGATCAAGTGGGAAACAGCTACAGCTCTGCAG 5420
5326 CGATGCTCTTGGGCTGCTGCTGGCCAAACGGCATGGAGGTGGCGCTCGAGCTGAGCC 5385
5421 CAACGGGACCTCGCGCTCATGTACGCCAAACGGCATGGGCGTCAGCTTCCACAGCGAGC 5480
5386 CCACTTGTGGCTGGCCACCGTCAACCCCAACCGTGGGCAAGAGAAATGTCAACGCTGCCAT 5445
5481 CCAGCTCTCGAGGACACCTTCAACCCCAACCTCGGCGCTGTAACTCTCCCTGCGCAT 5540
5446 CGACAAACCGCTCAACCTGGTGGAGTGGCGCAGCCAAAGAGCAGGCTCGGGGCGCAGGT 5505
5541 GGAGAAACCGCTGAACTCCATCGAGTGGCGCTGAGGAAGAAACAGATTAAGAGCAAGT 5600
5506 CACTGTCTTGGGCGCGCTCGGGTGCAACCGAAATCTCTATCTCTGACATTTGA 5565
5601 CACCATCTTGGAGGAAGCTCGGGTCCAGGAAGAAACCTCTCTGCTCATTTGATTATGA 5660
5566 TCGCTTAACACGACACAGAGAAGATCTATGATGACCAACCGCAAGTTTCAACCTTGGATTCT 5625
5661 CCGAAATATCCGACCTGAGAAGATCTATGACGACCAACCGGAAGTTTCAACCTGAGGATCAT 5720
5626 GTACGACAGGCGGGCGCGCCCTCTGGTCAACCCAGCAGCAGCTGAATGGTGTCAA 5685
5721 TTATGACAGGTGGGCGCGCTTCTGTGGTCTCCCGACGAGTGGAGTGGCGCGCTCAA 5780
5686 CGTGACATCTCCCTTGGGGTTTACATGCTGGCATCCAGAGGGCATCATGCTCTCAAAG 5745
5781 TGTCTCTTCTTCTTCAACGGGCGCTCGCGGCTTCCAGCGGGGCGCATGAGCGAG 5840
5746 AATGGAATACGACAGGCGGCGCATCATCCAGATCTTCTGTGATGGGAAGACATG 5805
5841 GACAGACATTTGACAAAGAGCGGATTTGTGCCGATTTTCCGCGAGCGGAAGTCTG 5900
5806 GAGCTACACATCTTAGAGAAGTCCATGCTGTGCTTACTACACAGCCAGGCGAGTAT 5865
5901 GAGCTATTCTTCTTGAAGTCCATGCTCTCTCTGCTGAGAGCAGCGCTGATCAT 5960
5866 CTTTGTAGTTCGAAAGATGACCGCTCTCTTCTGTGACGATGCCCAACGCTGCGCGGA 5925
5961 ATTTGAATATGACTCTCTGACCGCTTCCACGAGTCAACCTGCGCAGTGTGCGCGGGA 6020
5926 GACACTAGAGCACTCGGCTCAGTGGGCTACTACAGAACTCTATCACGCCCTGAGGG 5985

Qy	8134	CCAAGCGTGGCCCGCGAGCAGCAGAGACTGCGGGAAGGGAGGAGGCGCTGCGGGCGCTG	8193
Db	8220	TACTGCCCTGGGCCCAAGAGGACGACGAAAGCCAGGACGGAGAGAGGCGAGCCGTCTGTG	8279
Qy	8194	GACAGAGGGGAGAAAGCAGCAGGTGCTGAGCACAGGCGGGTGCAAGGCTACGACGGCTT	8253
Db	8280	GACGAGGGCGAGAAGCAGCAACTCCTGAGCACGGACGGGTGCAAGGTTATGAGGGCTA	8339
Qy	8254	TTTCGTGATCTCTGTCGAGCAGTACCGCAGAACTGTCAGACAGCGCCCAACATCCACTT	8313
Db	8340	TTACGTGCTTCCGGTGGACAGTACCCAGAGCTGGCAGACAGTAGCAGCAACATCCAGTT	8399
Qy	8314	CATGACACAGAGCGAGATGGCCCGGAGGTGACAGA	8348
Db	8400	CTTAAGACAGAAATGAGATGGGAAAGAGGTAAACAA	8434

Search completed: June 25, 2004, 03:28:00
 Job time : 31045 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 17:15:34 ; Search time 1916 Seconds
(without alignments)

18522.676 Million cell updates/sec

Title: US-10-029-020-13

Perfect score: 8354

Sequence: 1 gtttgatgtgaggagc.....ccggaggtgacagaggac 8354

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 337863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8354	100.0	8354	6	ABS52100 Human TEN
2	8225	98.5	8645	6	ABS78652 Human CDN
3	7779.2	93.1	8438	6	ABN85378 Human NOV
4	3504.6	42.0	8645	6	ABQ82344 Human NOV
5	3472.6	41.6	8675	6	ABQ82343 Human NOV
6	3424.4	41.0	8473	6	ABQ82345 Human NOV
7	3395.8	40.7	8487	6	ABQ82346 Human NOV
8	3385	40.5	9826	5	ASL14085 Human FCT
9	3362.6	40.3	9826	9	ADB32023 Human FCT
10	3351.4	40.1	9729	9	ASL14089 Human FCT
11	3351.4	40.1	9729	9	ADB32028 Human FCT
12	3274.2	39.2	13202	4	AAK51828 Human pol
13	3258	39.0	12879	6	ABK92230 Prostate
14	3165	37.9	9695	7	ACC72052 BC00205B
15	3068.6	36.7	9058	7	ACC72051 BC00205A
16	1708.8	20.5	3111	2	AAV19251 Human gam
17	1708.8	20.5	3111	2	AAK87705 Gamma-her
18	1708.8	20.5	3111	3	AAQ00791 Human Her
19	1708.8	20.5	3111	4	ASL18526 DNA encod
20	1562.6	18.7	4245	7	AAI60066 Human PCO
21	1560.4	18.7	3614	4	AAH14183 Human CDN
22	1437.4	17.2	3270	4	AAH14671 Human CDN
23	1193	14.3	3312	4	AAK52812 Human pol

ALIGNMENTS

RESULT 1

ABS52100

ID ABS52100 standard; DNA; 8354 BP.

XX AC ABS52100;

XX AC ABS52100;

DT DT 05-NOV-2002 (first entry)

XX DE Human TEN-M4-like gene.

XX DE Human TEN-M4-like gene.

Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis; cell signal processing; metabolic pathway modulation; metabolic disorder; obesity; diabetes; infectious disease; neurodegenerative disorder; acne; Alzheimer's disease; Parkinson's disease; immune disorder; cancer; haematopoietic disorder; cirrhosis; pancreatitis; learning defect; memory defect; infertility; congenital heart defect; hair growth; pigmentation disorder; endocrine disorder; respiratory disease; health; gastro-intestinal disease; reproductive; neurological disease; bone marrow transplantation; endocrine disease; allergy; inflammation; nephrological disorder; urinary system disorder; age-related disorder; neuropsychiatric disorder; EGF-related protein; SCUBEL; TEN-M4; adipocyte complement-related C1q tumour necrosis factor; out at first; beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter; type Ia membrane sushi-containing domain; butyrophilin; type Ia membrane-sushi domain containing; SNP; gene; ds; single nucleotide polymorphism.

OS Homo sapiens.

Key Location/Qualifiers

variation /tag= a

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= b

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= c

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= c

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= c

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= c

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= c

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= c

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= c

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= c

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= c

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= c

variation /standard_name= "Single nucleotide polymorphism (SNP)"

variation /tag= c

QY 1441 TCTGGAAAGGCAGCCCTGGTTGGCATTATATGACAGAAAGGCCTCCCTCCTTCAATAC 1500
Db 1441 TCTGGAAAGGCAGCCCTGGTTGGCATTATATGACAGAAAGGCCTCCCTCCTTCAATAC 1500
QY 1501 ACAGTTTGACTTGTGGAGCTGCTGGATGCGAGAGCTCCTAAACCCAGAGAGCGGGAG 1560
Db 1501 ACAGTTTGACTTGTGGAGCTGCTGGATGCGAGAGCTCCTAAACCCAGAGAGCGGGAG 1560
QY 1561 CCTAGAGGGACCCCGCCAGTCTCGGGAATGTGCCCCCTCCAGCCATCAGACAGG 1620
Db 1561 CCTAGAGGGACCCCGCCAGTCTCGGGAATGTGCCCCCTCCAGCCATCAGACAGG 1620
QY 1621 CTTTCATCCAGTATTGGATTAGGAATCTGGCACTTGGCTTTTACAATGACGGAAGGA 1680
Db 1621 CTTTCATCCAGTATTGGATTAGGAATCTGGCACTTGGCTTTTACAATGACGGAAGGA 1680
QY 1681 GTGAGAGTGGTTTCCCTTTCACCACTGCCATGAGTCGGTGAATACTGCCCGACAA 1740
Db 1681 GTGAGAGTGGTTTCCCTTTCACCACTGCCATGAGTCGGTGAATACTGCCCGACAA 1740
QY 1741 CTGCTATGGCAATGGTGAATCTCTGGCACTGCCACTGCTTCCCTGGGTTTCCCTGGG 1800
Db 1741 CTGCTATGGCAATGGTGAATCTCTGGCACTGCCACTGCTTCCCTGGGTTTCCCTGGG 1800
QY 1801 CCCGCACTGGGAGAGCTCTCCCGTGTCTGTAGCGGAATAGGCCAATACATGAA 1860
Db 1801 CCCGCACTGGGAGAGCTCTCCCGTGTCTGTAGCGGAATAGGCCAATACATGAA 1860
QY 1861 AGGAGATGCTTGTGCCACAGTGTGGAAGGCGTGAATGCGATGTCGCCACCAACCA 1920
Db 1861 AGGAGATGCTTGTGCCACAGTGTGGAAGGCGTGAATGCGATGTCGCCACCAACCA 1920
QY 1921 GTGTATCGATGTGGCCTGCAGCAACCATGGCACTGCGATCAAGGGCACTGCTATGCAA 1980
Db 1921 GTGTATCGATGTGGCCTGCAGCAACCATGGCACTGCGATCAAGGGCACTGCTATGCAA 1980
QY 1981 CCCTGGCTACAGGGCAGAGCTGTGAGGAAGTGGATGCGATGCGACCCCACTGTTGAGG 2040
Db 1981 CCCTGGCTACAGGGCAGAGCTGTGAGGAAGTGGATGCGATGCGACCCCACTGTTGAGG 2040
QY 2041 CCGGGGTGTCTGCGTGAGCGCAATGCAATGCTTTGTGGATGCGGAGCGCACTGCACTG 2100
Db 2041 CCGGGGTGTCTGCGTGAGCGCAATGCAATGCTTTGTGGATGCGGAGCGCACTGCACTG 2100
QY 2101 CGAGACCCCGAGGCGCAATGCTTAGACCAAGTGTTCAGGCGACCGAACTTCTCCCGGA 2160
Db 2101 CGAGACCCCGAGGCGCAATGCTTAGACCAAGTGTTCAGGCGACCGAACTTCTCCCGGA 2160
QY 2161 CACCGGCTTTCAGCTGTGACCAAGCTGGAAGTGGATGCGATGCTTCTATCGAGATCTG 2220
Db 2161 CACCGGCTTTCAGCTGTGACCAAGCTGGAAGTGGATGCGATGCTTCTATCGAGATCTG 2220
QY 2221 TGCTGCCGACTGTGGTGGCCATGGCGTGTGCTAGGGGGCACTTCCGCTGCGAGGATGG 2280
Db 2221 TGCTGCCGACTGTGGTGGCCATGGCGTGTGCTAGGGGGCACTTCCGCTGCGAGGATGG 2280
QY 2281 CTGGATGGGGCAGCTGCGACCGAGCGGCTGCGACCCCGCTGTGCGGAGATGGGAC 2340
Db 2281 CTGGATGGGGCAGCTGCGACCGAGCGGCTGCGACCCCGCTGTGCGGAGATGGGAC 2340
QY 2341 CTGCGCGACGCGCAAGTCCGAGTGCAGCCCTGGCTGGAATGGCGAAACACTGCAACATCGC 2400
Db 2341 CTGCGCGACGCGCAAGTCCGAGTGCAGCCCTGGCTGGAATGGCGAAACACTGCAACATCGC 2400
QY 2401 TCACATCTGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACCGGAG 2460
Db 2401 TCACATCTGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACCGGAG 2460
QY 2461 ATGTACCTTAGACCTGATGTTGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGG 2520
Db 2461 ATGTACCTTAGACCTGATGTTGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGG 2520
QY 2521 CTGTGACACTTCCATGGAGACTGCTGCGGTGACAGCAAGCAATGATGGAGATGGCCT 2580

Db 2521 CTGTGACACTTCCATGGAGACTGCTGCGGTGACAGCAAGCAATGATGAGATGGCCT 2580
QY 2581 GGTGACTGATGAGACCCCTGCTGCTCGAGCCCTGTGCCATATCAACCCGCTGTG 2640
Db 2581 GGTGACTGATGAGACCCCTGCTGCTCGAGCCCTGTGCCATATCAACCCGCTGTG 2640
QY 2641 CTTTGGCTCCCTTAACCCCTCTGGACATCATCCAGGAGACACAGGTCCCTGTGTCAAGCA 2700
Db 2641 CTTTGGCTCCCTTAACCCCTCTGGACATCATCCAGGAGACACAGGTCCCTGTGTCAAGCA 2700
QY 2701 GAACCTACATCTCTTATGACCGCATCAAGTTTCCTGTGGCGAGGACAGCACCAT 2760
Db 2701 GAACCTACATCTCTTATGACCGCATCAAGTTTCCTGTGGCGAGGACAGCACCAT 2760
QY 2761 AATCCCCGGGGAACCCCTTTGATGAGGAGCATCTTGTGTATTCCTGGCGCAAGTGT 2820
Db 2761 AATCCCCGGGGAACCCCTTTGATGAGGAGCATCTTGTGTATTCCTGGCGCAAGTGT 2820
QY 2821 GACATCAGATGGAACCCCTCTGGTTGTGTGAACATCATAGTTTGTCAATAACCCCTCT 2880
Db 2821 GACATCAGATGGAACCCCTCTGGTTGTGTGAACATCATAGTTTGTCAATAACCCCTCT 2880
QY 2881 TGGATATACATCAGCAGGCAAGATGGCAGCTTTGACTTGTGCAAAATGGCGCATCTC 2940
Db 2881 TGGATATACATCAGCAGGCAAGATGGCAGCTTTGACTTGTGCAAAATGGCGCATCTC 2940
QY 2941 CATCATCTCTGCGGTTTCGAGCGGCGCACCTTTTCATCACACAGGAGCACACCCCTGTG 3000
Db 2941 CATCATCTCTGCGGTTTCGAGCGGCGCACCTTTTCATCACACAGGAGCACACCCCTGTG 3000
QY 3001 ATGGGATGCTCTTTTGTTCATGGAACCATCATCATGACATGAGGAGAAATGAGATTCC 3060
Db 3001 ATGGGATGCTCTTTTGTTCATGGAACCATCATCATGACATGAGGAGAAATGAGATTCC 3060
QY 3061 CAGCTGTGACCTGAGCAATTTTGGCGCCCAACAGTCGCTCTCCATCCCCACTGAC 3120
Db 3061 CAGCTGTGACCTGAGCAATTTTGGCGCCCAACAGTCGCTCTCCATCCCCACTGAC 3120
QY 3121 GTCTTTCGCACTCCTGTGAGAAAGGCCCAATTTGTGCGGAAATTCAGGCTTTGCA 3180
Db 3121 GTCTTTCGCACTCCTGTGAGAAAGGCCCAATTTGTGCGGAAATTCAGGCTTTGCA 3180
QY 3181 GGAGGAAATCTATCTCTGCGCTGAAGATGAGGTGAGTACCTGAGCAGCGGACCCC 3240
Db 3181 GGAGGAAATCTATCTCTGCGCTGAAGATGAGGTGAGTACCTGAGCAGCGGACCCC 3240
QY 3241 TGGCTACAAATCTGCTCCTGAGGATCAGCCTCAACCCCGACCATCCCTTCAACCTCAT 3300
Db 3241 TGGCTACAAATCTGCTCCTGAGGATCAGCCTCAACCCCGACCATCCCTTCAACCTCAT 3300
QY 3301 GAAGGTGCACTCATGTTAGCGGTGGAGGCGGCTCTTCAGAAAGTGGTTGCGTGCAGC 3360
Db 3301 GAAGGTGCACTCATGTTAGCGGTGGAGGCGGCTCTTCAGAAAGTGGTTGCGTGCAGC 3360
QY 3361 CCAGACCTGTCTCTATTTTCAATTTGGGACAAAGACAGCTCTACAAAGAGGTGTT 3420
Db 3361 CCAGACCTGTCTCTATTTTCAATTTGGGACAAAGACAGCTCTACAAAGAGGTGTT 3420
QY 3421 TGGGCTTTTCAAGAGCCTTTTGGTGGGTTATGAATATGAATCCTGCGCCAGATCTAAT 3480
Db 3421 TGGGCTTTTCAAGAGCCTTTTGGTGGGTTATGAATATGAATCCTGCGCCAGATCTAAT 3480
QY 3481 CCTGTGGGAAAGAAAGCAACAGTGTGAGGCTATGAAATGACGCGTCCAGCTTGG 3540
Db 3481 CCTGTGGGAAAGAAAGCAACAGTGTGAGGCTATGAAATGACGCGTCCAGCTTGG 3540
QY 3541 AGGATGAGCCTAGACAAACATCATGCCCCTCAACATTCAAAGTGGTATCCTGCACAAAG 3600
Db 3541 AGGATGAGCCTAGACAAACATCATGCCCCTCAACATTCAAAGTGGTATCCTGCACAAAG 3600
QY 3601 GAAATGGGGAACCAAGTTTGTGTCTCAGAGCCTCTGTCTTTCATTTGGGAGCATCATGG 3660

Db	3601	GAATGGGAGAAACAGATTGTGTCTCAGACGCTCCTGTCAATTCGGAGCATCATGGGCAA	3660
Qy	3661	TGGGGCCGGAGAACATCTCTGCCCCAGCTGCAACGGCCCTTGCTGACGGCAACAGCT	3720
Db	3661	TGGGGCCGGAGAACATCTCTGCCCCAGCTGCAACGGCCCTTGCTGACGGCAACAGCT	3720
Qy	3721	CTGGCCCCAGTGGCCCTCACTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTCAA	3780
Db	3721	CTGGCCCCAGTGGCCCTCACTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTCAA	3780
Qy	3781	CTACATTAAGAGGATTTCCCTCTGGAAATGTCAACAAATCTTAGAGTGGAGATAA	3840
Db	3781	CTACATTAAGAGGATTTCCCTCTGGAAATGTCAACAAATCTTAGAGTGGAGATAA	3840
Qy	3841	AGATTTACAGATAGTACAGTCCAGCACAAATACTACTGGCCACAGACCCCATGAG	3900
Db	3841	AGATTTACAGATAGTACAGTCCAGCACAAATACTACTGGCCACAGACCCCATGAG	3900
Qy	3901	TGGGCCCTCTTCCCTTTCTGACAGCAACAGCCGGCGGTCTTTAAAAATCAAGTCCACTGT	3960
Db	3901	TGGGCCCTCTTCCCTTTCTGACAGCAACAGCCGGCGGTCTTTAAAAATCAAGTCCACTGT	3960
Qy	3961	GGTGGTAAGAACCTTGTCAAGAACTCTGAGTGGTTCGGGGACAGGTGACCAAGTGCCT	4020
Db	3961	GGTGGTAAGAACCTTGTCAAGAACTCTGAGTGGTTCGGGGACAGGTGACCAAGTGCCT	4020
Qy	4021	CCCTTTGATGACACTCGCTCGGGGATGGTGGGAAGGCCACAGAAGCCACACTCACCAA	4080
Db	4021	CCCTTTGATGACACTCGCTCGGGGATGGTGGGAAGGCCACAGAAGCCACACTCACCAA	4080
Qy	4081	TCCACAGGGATTAACAGTGGACAAGTTTGGGCTGATCTATTCTGTGATGGCACCATGAT	4140
Db	4081	TCCACAGGGATTAACAGTGGACAAGTTTGGGCTGATCTATTCTGTGATGGCACCATGAT	4140
Qy	4141	CAGACGATCATCAGATGGATCATCTCCACCTCTCGGCTCTTAATGATCTCACATC	4200
Db	4141	CAGACGATCATCAGATGGATCATCTCCACCTCTCGGCTCTTAATGATCTCACATC	4200
Qy	4201	AGCCCGGCCTCAGCTGTGATTTCTGATGGATTAATTCACAGTAAGACTGGAGTGGCC	4260
Db	4201	AGCCCGGCCTCAGCTGTGATTTCTGATGGATTAATTCACAGTAAGACTGGAGTGGCC	4260
Qy	4261	CACAGACTTACCATCACCAATGGACAACACTCACTTTATGTCTCGACAACATGTGGT	4320
Db	4261	CACAGACTTACCATCACCAATGGACAACACTCACTTTATGTCTCGACAACATGTGGT	4320
Qy	4321	CCTGCAATCTCTGAAACACCAAGTGGCATTTGTCGGGGAGGCCCATGCACTGCCA	4380
Db	4321	CCTGCAATCTCTGAAACACCAAGTGGCATTTGTCGGGGAGGCCCATGCACTGCCA	4380
Qy	4381	GGTCCCTGGATTAACCACTTCCTGTAAGCAAGTGGCCATCAACCAACCTGGAGTC	4440
Db	4381	GGTCCCTGGATTAACCACTTCCTGTAAGCAAGTGGCCATCAACCAACCTGGAGTC	4440
Qy	4441	AGCCACCGCTTGGCTGTTCACACAATGGGTCCTGTATATTGCTGAGACTGATGAGAA	4500
Db	4441	AGCCACCGCTTGGCTGTTCACACAATGGGTCCTGTATATTGCTGAGACTGATGAGAA	4500
Qy	4501	AAAGATCAACCGCATCAGGAGGTCAACACTAGTGGAGAGATCTCACTCGTCTGGGC	4560
Db	4501	AAAGATCAACCGCATCAGGAGGTCAACACTAGTGGAGAGATCTCACTCGTCTGGGC	4560
Qy	4561	CCCAGTGGCTGCTACATGTAATAATGATGCCACTGTGATTTTCTGGACCGCATGG	4620
Db	4561	CCCAGTGGCTGCTACATGTAATAATGATGCCACTGTGATTTTCTGGACCGCATGG	4620
Qy	4621	TTATGCCAAGGATCAAAAGTTAAATACCCCATCTTCTGGCTGTGTGCTGATGGGA	4680
Db	4621	TTATGCCAAGGATCAAAAGTTAAATACCCCATCTTCTGGCTGTGTGCTGATGGGA	4680
Qy	4681	GCTCTAGTGGCCACCTTGGGACATCCGAATTCGGTTTATCCGGAAGAACAGCCTTT	4740
Db	4681	GCTCTAGTGGCCACCTTGGGACATCCGAATTCGGTTTATCCGGAAGAACAGCCTTT	4740
Qy	4741	CCTCAACACCCAGAACATGTATGAGCTGTCTTCCAAATTTGACCAAGGAGCTCTATCTGTT	4800
Db	4741	CCTCAACACCCAGAACATGTATGAGCTGTCTTCCAAATTTGACCAAGGAGCTCTATCTGTT	4800
Qy	4801	TCATACCAACCCGCAAGCACTGTACACCAAAAGCCTGCCACAGGAGACTACTCTGTACAA	4860
Db	4801	TCATACCAACCCGCAAGCACTGTACACCAAAAGCCTGCCACAGGAGACTACTCTGTACAA	4860
Qy	4861	CTTCACTTACTGCGGAGCGGACATCACTATCATCAGACAAATATGGCAACATGGT	4920
Db	4861	CTTCACTTACTGCGGAGCGGACATCACTATCATCAGACAAATATGGCAACATGGT	4920
Qy	4921	AAATGTCTCGCGAGACTCTACTTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGT	4980
Db	4921	AAATGTCTCGCGAGACTCTACTTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGGT	4980
Qy	4981	GTACTGGGTGACCATGGGACCAACAGTGCACTCAAGAGTGTGACCAACAAAGCACAGA	5040
Db	4981	GTACTGGGTGACCATGGGACCAACAGTGCACTCAAGAGTGTGACCAACAAAGCACAGA	5040
Qy	5041	GTTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAA	5100
Db	5041	GTTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAA	5100
Qy	5101	CGGATGGGCAACATTTTATGATGACGACAGCTTTGGCGCTTGACAAATGTGACCTTCCC	5160
Db	5101	CGGATGGGCAACATTTTATGATGACGACAGCTTTGGCGCTTGACAAATGTGACCTTCCC	5160
Qy	5161	TACTGGCCAGGTGACAGCTTTCCGAGTGTACAGACAGTTTCCAGTGTGTCAGAGTAGA	5220
Db	5161	TACTGGCCAGGTGACAGCTTTCCGAGTGTACAGACAGTTTCCAGTGTGTCAGAGTAGA	5220
Qy	5221	GACCTCCAGCAAGGATGATGTCAACATAACCAACCACTGTCTGCTCAGCGGCTTCTA	5280
Db	5221	GACCTCCAGCAAGGATGATGTCAACATAACCAACCACTGTCTGCTCAGCGGCTTCTA	5280
Qy	5281	CACATGTGTGAAGACCAAGTCCGGAACAGCTATACATTCGGGCGCATGGCTCTTGCG	5340
Db	5281	CACATGTGTGAAGACCAAGTCCGGAACAGCTATACATTCGGGCGCATGGCTCTTGCG	5340
Qy	5341	GCTGTCTGTGGCCAAACGCAATGGAGTGGCGCTCGAGACTGAGCCCACTTGTGGCTGG	5400
Db	5341	GCTGTCTGTGGCCAAACGCAATGGAGTGGCGCTCGAGACTGAGCCCACTTGTGGCTGG	5400
Qy	5401	CACCGTCAACCCGACCGTGGGCAAGAGGAATGTCAACGTCGCCCATCGAACACGCTCAA	5460
Db	5401	CACCGTCAACCCGACCGTGGGCAAGAGGAATGTCAACGTCGCCCATCGAACACGCTCAA	5460
Qy	5461	CCTGTGTGAGTGGCCGACCGAAGAGCAGCTCGGGGCCAGGTCTACTGTCTTTGGCG	5520
Db	5461	CCTGTGTGAGTGGCCGACCGAAGAGCAGCTCGGGGCCAGGTCTACTGTCTTTGGCG	5520
Qy	5521	CCGGCTGCGGGTGCACAAACCGAAATCTCTGATCTTCTGATCGCGTAAACACGCAAC	5580
Db	5521	CCGGCTGCGGGTGCACAAACCGAAATCTCTGATCTTCTGATCGCGTAAACACGCAAC	5580
Qy	5581	AGAGAAGATCTATGATGACCAACCGAAGTTCACCTTTCGGATTCTGTACGACGAGGGG	5640
Db	5581	AGAGAAGATCTATGATGACCAACCGAAGTTCACCTTTCGGATTCTGTACGACGAGGGG	5640
Qy	5641	CGGGCCAGCTCTGGTTCACCCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCC	5700
Db	5641	CGGGCCAGCTCTGGTTCACCCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCC	5700
Qy	5701	TGGGGGTTTACATGCTGGCATCCAGAGGGCATCATGTCTGAAAGAAATGGAATACGACCA	5760
Db	5701	TGGGGGTTTACATGCTGGCATCCAGAGGGCATCATGTCTGAAAGAAATGGAATACGACCA	5760
Qy	5761	GGCGGGCCGATCATCCAGGATCTTGGTGTGGGAAGACATGGAGCTACACATCTT	5820
Db	5761	GGCGGGCCGATCATCCAGGATCTTGGTGTGGGAAGACATGGAGCTACACATCTT	5820

QY	5821	AGAGAGTCCATGTGTGCTGCTACTACAGCAGCAGGCGAGTATATCTTTGAGTTGACAA	5880	Db	6901	CAGGTACCGTACGATGCGCTGGGCGGCGGTGTCCAGCAAGCAGCCACAGCCACA	6960
Db	5821	AGAGAGTCCATGTGTGCTGCTACTACAGCAGCAGGCGAGTATATCTTTGAGTTGACAA	5880	QY	6961	CTGCAGATTCTTCTATGAGACCTGACCAACCCCAAGGTCAACCACTGCTGACCA	7020
QY	5881	GAATGACCGCTCTCTTCTGAGCATGCCCCAGCTGCGCGGCGAGACACTAGAGACCAT	5940	Db	6961	CTGCAGATTCTTCTATGAGACCTGACCAACCCCAAGGTCAACCACTGCTGACCA	7020
Db	5881	GAATGACCGCTCTCTTCTGAGCATGCCCCAGCTGCGCGGCGAGACACTAGAGACCAT	5940	QY	7021	CTCCAGCTCTGAGATCACCTCCCTCTACTAGCACTTGAAGGACACCTCTTTGGCATGA	7080
QY	5941	CGCTCAGTGGCTTACTACAGAAACATCTATCAGCCGCTGAGGGCAATGCCCTCAGTCAT	6000	Db	7021	CTCCAGCTCTGAGATCACCTCCCTCTACTAGCACTTGAAGGACACCTCTTTGGCATGA	7080
Db	5941	CGCTCAGTGGCTTACTACAGAAACATCTATCAGCCGCTGAGGGCAATGCCCTCAGTCAT	6000	QY	7081	GCTGAGCAGTGTGTGATGAGTTTATAGCTTTGTGACCAACATCGGGACCCCTCTTCTGT	7140
QY	6001	ACAGGACTTCACTGAGGATGGGCACTCTTACACACCTTCTACCTGGGCACTGGCGCAG	6060	Db	7081	GCTGAGCAGTGTGTGATGAGTTTATAGCTTTGTGACCAACATCGGGACCCCTCTTCTGT	7140
Db	6001	ACAGGACTTCACTGAGGATGGGCACTCTTACACACCTTCTACCTGGGCACTGGCGCAG	6060	QY	7141	CTTTAGTGAAACAGGTTTGTATGATCAAGCAAACTCTGTACAGGCTATGGGAGATCTA	7200
QY	6061	GCTGATATCAAGTATGGCAAACTGTCAAAGCTGGGAGAGCGCTCTATGACACCA	6120	Db	7141	CTTTAGTGAAACAGGTTTGTATGATCAAGCAAACTCTGTACAGGCTATGGGAGATCTA	7200
Db	6061	GCTGATATCAAGTATGGCAAACTGTCAAAGCTGGGAGAGCGCTCTATGACACCA	6120	QY	7201	CATGATATCAACCCCAACTTTTTCAGATCATATAGGCTACCATGGTGGCTCTATGATCC	7260
QY	6121	GCTCAGTTTCACTATGACGAGAGCGGAGGATGCTGAAGACCACTCAACCTACAGAAATGA	6180	Db	7201	CATGATATCAACCCCAACTTTTTCAGATCATATAGGCTACCATGGTGGCTCTATGATCC	7260
Db	6121	GCTCAGTTTCACTATGACGAGAGCGGAGGATGCTGAAGACCACTCAACCTACAGAAATGA	6180	QY	7261	ACTCACCAGCTTGTCCACATGGGCGGCGGAGATTATGATGTCTGGCGGAGCGTGAC	7320
QY	6181	GGGCTTCACTGACAGTCCGCTACCGTACAGTTGGGCGGCTGATGACCGACAGATCTT	6240	Db	7261	ACTCACCAGCTTGTCCACATGGGCGGCGGAGATTATGATGTCTGGCGGAGCGTGAC	7320
Db	6181	GGGCTTCACTGACAGTCCGCTACCGTACAGTTGGGCGGCTGATGACCGACAGATCTT	6240	QY	7321	TAGCCAGACACGAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCAATGCTTTTAACT	7380
QY	6241	CCGCTTCACTGAGGAGCATGCTCAAGCCGCTTTGACTACAACTATGACACAGCTT	6300	Db	7321	TAGCCAGACACGAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCAATGCTTTTAACT	7380
Db	6241	CCGCTTCACTGAGGAGCATGCTCAAGCCGCTTTGACTACAACTATGACACAGCTT	6300	QY	7381	CTATATGTTTCAAAAACAAACCCCATCAGCAACTCCCAGGACATCAAGTGTTCATGAC	7440
QY	6301	CCGCGTGACAGGCTGTGATCAAGAGACCCCACTGCGCCATTTGATCTCTATCG	6360	Db	7381	CTATATGTTTCAAAAACAAACCCCATCAGCAACTCCCAGGACATCAAGTGTTCATGAC	7440
Db	6301	CCGCGTGACAGGCTGTGATCAAGAGACCCCACTGCGCCATTTGATCTCTATCG	6360	QY	7441	AGATGTTTAAAGCTGCTCCTCAGCTTTCAGCTTACCAAGCTCATCCCTGGTTA	7500
QY	6361	CTATGATGATGTGACGAGACAGAGCAGTTTGGGAGTTTGGTCTCAATTACTATGA	6420	Db	7441	AGATGTTTAAAGCTGCTCCTCAGCTTTCAGCTTACCAAGCTCATCCCTGGTTA	7500
Db	6361	CTATGATGATGTGACGAGACAGAGCAGTTTGGGAGTTTGGTCTCAATTACTATGA	6420	QY	7501	TCCCAAACAGACATGATGCCATGGAACCCCTCCTACGAGCTCATCCACACAGATGAA	7560
QY	6421	CATTAAACAGATCATCACCAGCTGTATCAGCCACACCAAGCATTTGATGATATGG	6480	Db	7501	TCCCAAACAGACATGATGCCATGGAACCCCTCCTACGAGCTCATCCACACAGATGAA	7560
Db	6421	CATTAAACAGATCATCACCAGCTGTATCAGCCACACCAAGCATTTGATGATATGG	6480	QY	7561	AACGAGGAGTGGGACACAGCAAGTCTATCTCCGGGTACAGTGTGAAGTACAGAGCA	7620
QY	6481	CAGATGAAGGAGTGCAGTATGAGATCTCCGCTGCTGATGCTGATGAGTGCCTCA	6540	Db	7561	AACGAGGAGTGGGACACAGCAAGTCTATCTCCGGGTACAGTGTGAAGTACAGAGCA	7620
Db	6481	CAGATGAAGGAGTGCAGTATGAGATCTCCGCTGCTGATGCTGATGAGTGCCTCA	6540	QY	7621	GCTCAAGGCTTTGTACCTTAGAACGGTTTGACCAAGCTCTATGGCTCCACCAATCACCAG	7680
QY	6541	GTATGATAACATGGGCGAGTAGTGAAGAGAGTGAAGAGTGAAGTGAAGTGAAGTGA	6600	Db	7621	GCTCAAGGCTTTGTACCTTAGAACGGTTTGACCAAGCTCTATGGCTCCACCAATCACCAG	7680
Db	6541	GTATGATAACATGGGCGAGTAGTGAAGAGAGTGAAGAGTGAAGTGAAGTGAAGTGA	6600	QY	7681	CTGCCAGAGGCTCCAAAGACCAAGATTTGCAATGAGTTCAGCGCTCAGTCTTTGGCAAGG	7740
QY	6601	CACCTGCTACTCTATGATGATGATGCTGACGCGGCTGATGATGATGATGATGATGAT	6660	Db	7681	CTGCCAGAGGCTCCAAAGACCAAGATTTGCAATGAGTTCAGCGCTCAGTCTTTGGCAAGG	7740
Db	6601	CACCTGCTACTCTATGATGATGATGCTGACGCGGCTGATGATGATGATGATGATGAT	6660	QY	7741	GGTCAAGTTTGTCTTGAAGGATGGCGAGTGGACCAAGATCATCATGAGTGTGGCCAAATGA	7800
QY	6661	CAAGCCACTCTGGGCTACAGTACGCTCAATGGGACCTGAGGACCTGAGGACCTGAG	6720	Db	7741	GGTCAAGTTTGTCTTGAAGGATGGCGAGTGGACCAAGATCATCATGAGTGTGGCCAAATGA	7800
Db	6661	CAAGCCACTCTGGGCTACAGTACGCTCAATGGGACCTGAGGACCTGAGGACCTGAG	6720	QY	7801	GGATGGCGAAGGTTGCTGCCATCTTGAACCATGATGCCACTACCTAGAGAACTGACCTT	7860
QY	6721	GAAAGTGCAGGCTTCAACCACTACCGTATGATGATGATGATGATGATGATGATGATGAT	6780	Db	7801	GGATGGCGAAGGTTGCTGCCATCTTGAACCATGATGCCACTACCTAGAGAACTGACCTT	7860
Db	6721	GAAAGTGCAGGCTTCAACCACTACCGTATGATGATGATGATGATGATGATGATGATGAT	6780	QY	7861	CACCATTCCTCAGTGGGGCGGCGGAGAACTTGTGAAACCCAGGACCTTCAAGAGGTGACCT	7920
QY	6781	TGACGTGCAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	6840	Db	7861	CACCATTCCTCAGTGGGGCGGCGGAGAACTTGTGAAACCCAGGACCTTCAAGAGGTGACCT	7920
Db	6781	TGACGTGCAATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	6840	QY	7921	GGCCATTCCTCAGTGGGGCGGCGGAGAACTTGTGAAACCCAGGACCTTCAAGAGGTGACCT	7980
QY	6841	TGAGTACAACTCAGCTGCTTCAATCAAGGCTTCAACCGGCTGAGGAGTGGAGTGG	6900	Db	7921	GGCCATTCCTCAGTGGGGCGGCGGAGAACTTGTGAAACCCAGGACCTTCAAGAGGTGACCT	7980
Db	6841	TGAGTACAACTCAGCTGCTTCAATCAAGGCTTCAACCGGCTGAGGAGTGGAGTGG	6900	QY	7981	GTCCAGATCAACAGTACTTAAAGGAGGAGTGAAGCTTACAGACATCCAGCTCCA	8040
QY	6901	CAGGTACCGCTACGATGGCTGGGCGGCGGCTGTCCAGCAAGAGCAGCCACAGCCACA	6960				

QY 481 TTCCAATCTCACATCACCGACACCGAGCATGAATAACACTGAGACTGATCATCTCGGGCGG 540
Db 563 TTCCAATCTCACATCACCGACACCGAGCATGAATAACACTGAGAC-----TCCGGCGG 616
QY 541 CTTGAGAAACACGCGCGGCTCCGGAACGCGCGCGCGCGCTCTCGCAAGCCCAACACCC 600
Db 617 CTTGAGAAACACGCGCGGCTCCGGAACGCGCGCGCGCGCTCTCGCAAGCCCAACACCC 676
QY 601 CAACGAGACACGCGCGGCTCCATTAATCTCCCTGAACCGGGGCACTTTCAGCGCGAGGAG 660
Db 677 CAACGAGACACGCGCGGCTCCATTAATCTCCCTGAACCGGGGCACTTTCAGCGCGAGGAG 736
QY 661 CAACCGGACGCGCGGCTCCGGAACGCGCGCGCGCGCTCTCGCAAGCCCAACACCC 720
Db 737 CAACCGGACGCGCGGCTCCGGAACGCGCGCGCGCGCTCTCGCAAGCCCAACACCC 796
QY 721 CCAGGAGCTCCGCGCGGCTCCATTAATCTCCCTGAACCGGGGCACTTTCAGCGCGAGGAG 780
Db 797 CCAGGAGCTCCGCGCGGCTCCATTAATCTCCCTGAACCGGGGCACTTTCAGCGCGAGGAG 856
QY 781 CAGGAACCTAGGCAAGCAGGCTTCTTAGGACCATTTAGGACCATTTAGGACCATTTAGGATGA 840
Db 857 CAGGAACCTAGGCAAGCAGGCTTCTTAGGACCATTTAGGACCATTTAGGACCATTTAGGATGA 916
QY 841 CATTTCTCGCGGCTCCGCGGCTTCTTAGGACCATTTAGGACCATTTAGGACCATTTAGGATGA 900
Db 917 CATTTCTCGCGGCTCCGCGGCTTCTTAGGACCATTTAGGACCATTTAGGACCATTTAGGATGA 976
QY 901 TGGAGGACCTCCGCGGCTTCTTAGGACCATTTAGGACCATTTAGGACCATTTAGGATGA 960
Db 977 TGGAGGACCTCCGCGGCTTCTTAGGACCATTTAGGACCATTTAGGACCATTTAGGATGA 1036
QY 961 AGTGTACTCTCTCGCGGCTCCGCGGCTTCTTAGGACCATTTAGGACCATTTAGGATGA 1020
Db 1037 AGTGTACTCTCTCGCGGCTCCGCGGCTTCTTAGGACCATTTAGGACCATTTAGGATGA 1096
QY 1021 CCTAAGAAAGCCCTCCAAAGTACTGTAACTGGAAGTGGCAGCTTGGAGCGCATTCGTCAT 1080
Db 1097 CCTAAGAAAGCCCTCCAAAGTACTGTAACTGGAAGTGGCAGCTTGGAGCGCATTCGTCAT 1156
QY 1081 CTCAGCCACTCTGTGTCATCTCTGTCGATATCTTTGTGGCCATGACCTTTGTGCTTAA 1140
Db 1157 CTCAGCCACTCTGTGTCATCTCTGTCGATATCTTTGTGGCCATGACCTTTGTGCTTAA 1210
QY 1141 CTGGCACTCTGAGCGGCTGAGGCGGAGATGATGAGATCAGGAGGACACGCGGAG 1200
Db 1211 CTGGCACTCTGAGCGGCTGAGGCGGAGATGATGAGATCAGGAGGACACGCGGAG 1270
QY 1201 TTGGCTGTGCCAACCGACGCTCTCCCTATACCCCTCAGGGGGCACTGGCTTAGAGACCC 1260
Db 1271 TTGGCTGTGCCAACCGACGCTCTCCCTATACCCCTCAGGGGGCACTGGCTTAGAGACCC 1330
QY 1261 TGACGGAAGGCAAGGAAACACAGAAAGGAAAGCCAGTAGTTCTTTCCAGAGGACAG 1320
Db 1331 TGACGGAAGGCAAGGAAACACAGAAAGGAAAGCCAGTAGTTCTTTCCAGAGGACAG 1390
QY 1321 TTTCTAGATTTCTGAGAAATTTGATGTGGAAAGGAGCGCTCCAGAAAGATTCTCTCTGG 1380
Db 1391 TTTCTAGATTTCTGAGAAATTTGATGTGGAAAGGAGCGCTTTCCAGAAAGATTCTCTCTGG 1450
QY 1381 CACTTTCTGAGATTTCTAAGTGTTCATAGACCATCTCTGTGATCTGAAATTCATATGTC 1440
Db 1451 CACTTTCTGAGATTTCTAAGTGTTCATAGACCATCTCTGTGATCTGAAATTCATATGTC 1510
QY 1441 TCTGGAAAGGCAAGGCTTGTGATTTATGGAAGGAAAGGCTCTCTCTCTCAGATAC 1500
Db 1511 TCTGGAAAGGCAAGGCTTGTGATTTATGGAAGGAAAGGCTCTCTCTCTCAGATAC 1570
QY 1501 ACAGTTTCACTTTGTGAGCTGCTGATGGAGGAGCTCTCTAACCAGGAGGCGCGAG 1560
Db 1571 ACAGTTTCACTTTGTGAGCTGCTGATGGAGGAGCTCTCTAACCAGGAGGCGCGAG 1630

QY 1561 CCTAGAGGGGACCCCGCGCAGTCTCGGGGAACTGTGCCCCCTCCAGGCATGAGACAGG 1620
Db 1631 CCTAGAGGGGACCCCGCGCAGTCTCGGGGAACTGTGCCCCCTCCAGGCATGAGACAGG 1690
QY 1621 CTTTCATCCAGTATTGGATTCAGGAATCTGGCACTTGGCTTTTACAATGACGGAAGGA 1680
Db 1691 CTTTCATCCAGTATTGGATTCAGGAATCTGGCACTTGGCTTTTACAATGACGGAAGGA 1750
QY 1681 GTCAGAAGTGGTTTCTTCTTCTCACCACCTGCTGATGGTGGTGAATCTGCCCCAGCAA 1740
Db 1751 GTCAGAAGTGGTTTCTTCTTCTCACCACCTGCTGATGGTGGTGAATCTGCCCCAGCAA 1810
QY 1741 CTCTATGGCAATGGTGTGATCTTGGACCTGCCACTGCTTCTGGTTTCTGGG 1800
Db 1811 CTCTATGGCAATGGTGTGATCTTGGACCTGCCACTGCTTCTGGTTTCTGGG 1870
QY 1801 CCCCGACTGTGGCAGAGCTCTCTGCCCTGCTGTAGCGGAATGGCAATACATGAA 1860
Db 1871 CCCCGACTGTGGCAGAGCTCTCTGCCCTGCTGTAGCGGAATGGCAATACATGAA 1930
QY 1861 AGCAGATGCTTGTGCCACAGTGGCTGGAAGCGCTGAGTGGCATGTGCCCAACCA 1920
Db 1931 AGCAGATGCTTGTGCCACAGTGGCTGGAAGCGCTGAGTGGCATGTGCCCAACCA 1990
QY 1921 GTCTATCGATGTGGCTGAGCAACCATGGCACTGCTGATCAGCGGCACCTGATCTGCAA 1980
Db 1991 GTCTATCGATGTGGCTGAGCAACCATGGCACTGCTGATCAGCGGCACCTGATCTGCAA 2050
QY 1981 CCTTGGCTCAAGGGGCGAGAGTGTGAGGAAGTGGACTGCTGAGCCCAACATGTTTCAGG 2040
Db 2051 CCTTGGCTCAAGGGGCGAGAGTGTGAGGAAGTGGACTGCTGAGCCCAACATGTTTCAGG 2110
QY 2041 CCGGGTGTCTGCTGAGAGGCGAATGCCAATGCTTTTGGGATGGGGAGGACCAACTG 2100
Db 2111 CCGGGTGTCTGCTGAGAGGCGAATGCCAATGCTTTTGGGATGGGGAGGACCAACTG 2170
QY 2101 CGAGACCCCGAGGCGCAATGTTAGACCAAGTGGTTCAGGCCACGGAACCTTCTCCCGGA 2160
Db 2171 CGAGACCCCGAGGCGCAATGTTAGACCAAGTGGTTCAGGCCACGGAACCTTCTCCCGGA 2230
QY 2161 CACCGGCTTTCAGCTGTGACCCAGCTGGAACGCTGGAACGCTGTTTCTATCGAGATCTG 2220
Db 2231 CACCGGCTTTCAGCTGTGACCCAGCTGGAACGCTGGAACGCTGTTTCTATCGAGATCTG 2290
QY 2221 TGTGTCGAGCTGTGTGGCCATGCGGTGTGCTAGGGGCACTGCGCTGCGGAGATG 2280
Db 2291 TGTGTCGAGCTGTGTGGCCATGCGGTGTGCTAGGGGCACTGCGCTGCGGAGATG 2350
QY 2281 CTGGATGGGGGAGCCTGCGACAGCGGCTGCGCACCCGCGCTGTGCGGAGCATGGAC 2340
Db 2351 CTGGATGGGGGAGCCTGCGACAGCGGCTGCGCACCCGCGCTGTGCGGAGCATGGAC 2410
QY 2341 CTGCGCGGACGGCAAGTGGAGTGCAGCCCTGGCTGGAAATGGGAAACACTGCAACCATGCG 2400
Db 2411 CTGCGCGGACGGCAAGTGGAGTGCAGCCCTGGCTGGAAATGGGAAACACTGCAACCATGCG 2470
QY 2401 TCACTATCTGGATAGGATAGTTAAAGAGGTTGCCCTGGTTGTGCAATGGCAACGGCAG 2460
Db 2471 TCACTATCTGGATAGGATAGTTAAAGAGGTTGCCCTGGTTGTGCAATGGCAACGGCAG 2530
QY 2461 ATGTACTTGTAGCTGAATGGTGGCACTGCTGCTGCCAGCTGGCTGGAGAGGAGCTGG 2520
Db 2531 ATGTACTTGTAGCTGAATGGTGGCACTGCTGCTGCCAGCTGGCTGGAGAGGAGCTGG 2590
QY 2521 CTGTGACACTTCCATGGAGCTGCGCTGCGTGCAGCAAGCAATGATGGAGATGGCCT 2580
Db 2591 CTGTGACACTTCCATGGAGCTGCGCTGCGTGCAGCAAGCAATGATGGAGATGGCCT 2650
QY 2581 GGTGGACTGCATGGACCTTGAATGTGCTTCCAGCCCTGTGCCATATCAACCCGCTGTG 2640
Db 2651 GGTGGACTGCATGGACCTTGAATGTGCTTCCAGCCCTGTGCCATATCAACCCGCTGTG 2710
QY 2641 CCTTGGCTCCCTAACCCCTCTGGACATCATCCAGGAGACAGGTCCTCTGTGTACAGCA 2700

QY	4861	CTTACCTACACTGGGAGCGGACATCATCATCATACAGACAAATGGCAATGGT	4920
Db	4910	CTTACCTACACTGGGAGCGGACATCATCATCATACAGACAAATGGCAATGGT	4969
QY	4921	AAATGTCCGCGAGACTCTACTGGGATCCCTCTGGCTGGTGGTCCAGATGGCCAGT	4980
Db	4970	AAATGTCCGCGAGACTCTACTGGGATCCCTCTGGCTGGTGGTCCAGATGGCCAGT	5029
QY	4981	GTACTGGGTGACATGGGACCAACAGTGCACTCAAGAGTGTGACCAACAGGACAGGA	5040
Db	5030	GTACTGGGTGACATGGGACCAACAGTGCACTCAAGAGTGTGACCAACAGGACAGGA	5089
QY	5041	GTGGCCCATGATGACATACCATGGCAATTCGGGCTCTCTGGCAACCAAAAGCAATGAAA	5100
Db	5090	GTGGCCCATGATGACATACCATGGCAATTCGGGCTCTCTGGCAACCAAAAGCAATGAAA	5149
QY	5101	CGGATGGGACAACTTTATGATGACAGAGCTTTGGCGCCCTGACAAATGTGACTTCCC	5160
Db	5150	CGGATGGGACAACTTTATGATGACAGAGCTTTGGCGCCCTGACAAATGTGACTTCCC	5209
QY	5161	TACTGGCCAGGTGAGCAGTTTCCGAAGTATACAGACAGTTTCAGTGCAATGTCAGGTAGA	5220
Db	5210	TACTGGCCAGGTGAGCAGTTTCCGAAGTATACAGACAGTTTCAGTGCAATGTCAGGTAGA	5269
QY	5221	GACTCCAGCAAGGATGTACCAATACCAACCAACCTGTCTGCCTCAGCGCCCTTCTA	5280
Db	5270	GACTCCAGCAAGGATGTACCAATACCAACCAACCTGTCTGCCTCAGCGCCCTTCTA	5329
QY	5281	CACACTGTGCAAGACCAAGTCCGGAACAGTACTACATCCGGGGCCGATGGCTCTTGGC	5340
Db	5330	CACACTGTGCAAGACCAAGTCCGGAACAGTACTACATCCGGGGCCGATGGCTCTTGGC	5389
QY	5341	GCTGTGTGSCCAACGGCATGGAGTGGCGCTGCAGACTGAGCCCACTGTGCTGGCTGG	5400
Db	5390	GCTGTGTGSCCAACGGCATGGAGTGGCGCTGCAGACTGAGCCCACTGTGCTGGCTGG	5449
QY	5401	CACCGTCAACCCACCGTGGGACAGAGGAATGTACGCTGCCATTCGACAAACGGCTCAA	5460
Db	5450	CACCGTCAACCCACCGTGGGACAGAGGAATGTACGCTGCCATTCGACAAACGGCTCAA	5509
QY	5461	CCTGTGTGAGTGGGCGCCAGCGCAAGAGCAGGCTCGGGCCAGCTCACTGCTTTTGGCG	5520
Db	5510	CCTGTGTGAGTGGGCGCCAGCGCAAGAGCAGGCTCGGGCCAGCTCACTGCTTTTGGCG	5569
QY	5521	CCGGTGGCGGTGACAAACGAAATCTCTATCTCTGCACTTTGATCGCGTAAACAGCAC	5580
Db	5570	CCGGTGGCGGTGACAAACGAAATCTCTATCTCTGCACTTTGATCGCGTAAACAGCAC	5629
QY	5581	AGAGAAGATCTATGATGACCAACCGCAAGTTCACCCCTTCGGATTCGTACGACAGGCGG	5640
Db	5630	AGAGAAGATCTATGATGACCAACCGCAAGTTCACCCCTTCGGATTCGTACGACAGGCGG	5689
QY	5641	CGGGCCAGCCCTCTGGTCAACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATCTCCCC	5700
Db	5690	CGGGCCAGCCCTCTGGTCAACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATCTCCCC	5749
QY	5701	TGGGGTTACATTTGCTGATCCAGAGGGGATCATGTCTGAAGAATGGAATACGACCA	5760
Db	5750	TGGGGTTACATTTGCTGATCCAGAGGGGATCATGTCTGAAGAATGGAATACGACCA	5809
QY	5761	GGCGGGCCGATCATCATCCAGATCTTCGCTGATGGGAAGACATGGAGCTACACATCTT	5820
Db	5810	GGCGGGCCGATCATCATCCAGATCTTCGCTGATGGGAAGACATGGAGCTACACATCTT	5869
QY	5821	AGAGAAGTCCATTTGCTGCTTACACAGCCAGGACGATATCTTTTCAAGTTCGACAA	5880
Db	5870	AGAGAAGTCCATTTGCTGCTTACACAGCCAGGACGATATCTTTTCAAGTTCGACAA	5929
QY	5881	GAATGACCGCTCTCTTCTGTGAGATGCCCAAGTCCGGCGGACAGACTAGAGCCAT	5940
Db	5930	GAATGACCGCTCTCTTCTGTGAGATGCCCAAGTCCGGCGGACAGACTAGAGCCAT	5989
QY	5941	CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTCAGGGCAATGCCTCAGTCAT	6000
Db	5990	CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTCAGGGCAATGCCTCAGTCAT	6049
QY	6001	ACAGGACTTCACTGAGGATGGGCACCTCCTTACACCTTCTACTCTGGGCACTGGCGCAG	6060
Db	6050	ACAGGACTTCACTGAGGATGGGCACCTCCTTACACCTTCTACTCTGGGCACTGGCGCAG	6109
QY	6061	GGTGATATACAAGTATGGGMAACTGTCAAAGCTGGGAGAGACGCTCTATGACACCAAA	6120
Db	6110	GGTGATATACAAGTATGGGMAACTGTCAAAGCTGGGAGAGACGCTCTATGACACCAAA	6169
QY	6121	GGTCAGTTTCACTATGACGAGACGCGCATGTCTGAAGACCATCAACTCAGAGATGA	6180
Db	6170	GGTCAGTTTCACTATGACGAGACGCGCATGTCTGAAGACCATCAACTCAGAGATGA	6229
QY	6181	GGGCTTCACTGCAACCATCCGTACGATTTGGGCCCCCTGATTTGACCGACAGATCTT	6240
Db	6230	GGGCTTCACTGCAACCATCCGTACGATTTGGGCCCCCTGATTTGACCGACAGATCTT	6289
QY	6241	CCGCTTCACTGAGGAAGGATGGTCAACCGCCGTTTGGACTACAACTATGACAAACAGCTT	6300
Db	6290	CCGCTTCACTGAGGAAGGATGGTCAACCGCCGTTTGGACTACAACTATGACAAACAGCTT	6349
QY	6301	CCGGGTGACAGCATGACGCTGTGATCAACGAGACCCCACTGCCCATTTGATCTATCG	6360
Db	6350	CCGGGTGACAGCATGACGCTGTGATCAACGAGACCCCACTGCCCATTTGATCTATCG	6409
QY	6361	CTATGATGATGTCTCAGGCAAGACAGAGAGTTTGGAAAGTTTGGTGTCTATTACTATGA	6420
Db	6410	CTATGATGATGTCTCAGGCAAGACAGAGAGTTTGGAAAGTTTGGTGTCTATTACTATGA	6469
QY	6421	CATTAAACAGATCATCACCACAGCTGTCTATGATCCCAACCAAGCATTTTGTGATATGG	6480
Db	6470	CATTAAACAGATCATCACCACAGCTGTCTATGATCCCAACCAAGCATTTTGTGATATGG	6529
QY	6481	CAGGATGAAGGAAGTCCAGTATGAGATCTTCCGCTCGCTCATGTACTGTGATGACCTCCA	6540
Db	6530	CAGGATGAAGGAAGTCCAGTATGAGATCTTCCGCTCGCTCATGTACTGTGATGACCTCCA	6589
QY	6541	GTATGATTAACATGGGCGGAGTAGTGAAGAGGAGCTTGAAGGTAGGAGCCCTAGCCCAATAC	6600
Db	6590	GTATGATTAACATGGGCGGAGTAGTGAAGAGGAGCTTGAAGGTAGGAGCCCTAGCCCAATAC	6649
QY	6601	CACCTCCTACTCCTATGATGATGATGCTGACGCGCCAGCTGACAGACAGTCTCCATCAATGA	6660
Db	6650	CACCTCCTACTCCTATGATGATGATGCTGACGCGCCAGCTGACAGACAGTCTCCATCAATGA	6709
QY	6661	CAAGCCACTCTGGCGCTCAGCTACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGG	6720
Db	6710	CAAGCCACTCTGGCGCTCAGCTACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGG	6769
QY	6721	GAACAGTGCACGGCTCACACCACTACGGTATGACATCCCGGACCGCATCACTCGGCTGG	6780
Db	6770	GAACAGTGCACGGCTCACACCACTACGGTATGACATCCCGGACCGCATCACTCGGCTGG	6829
QY	6781	TGACGTGCAATACAAGTGGATGGCTTCTGAGGACAGCGGGCGGTGATATCTT	6840
Db	6830	TGACGTGCAATACAAGTGGATGGCTTCTGAGGACAGCGGGCGGTGATATCTT	6889
QY	6841	TGAGTACAACTCAGCTGGCTCTCATCAAGGCTTACAAACCGGCTGGCAGCTGGAGTGT	6900
Db	6890	TGAGTACAACTCAGCTGGCTCTCATCAAGGCTTACAAACCGGCTGGCAGCTGGAGTGT	6949
QY	6901	CAGGTAACCGCTACGATGGGCTGGGCGCGCTGTCTCCAGCAAGACAGCAGCAGCACCA	6960
Db	6950	CAGGTAACCGCTACGATGGGCTGGGCGCGCTGTCTCCAGCAAGACAGCAGCAGCACCA	7009
QY	6961	CCTGCACTTCTTCTATGAGACCTGACCAACCCCAACCAAGTCAACCCCTGTACCAACA	7020
Db	7010	CCTGCACTTCTTCTATGAGACCTGACCAACCCCAACCAAGTCAACCCCTGTACCAACA	7069
QY	7021	CTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGAAGGACACCTCTTTTGCATGGA	7080

Db	7070	CTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGAAGGACACCTCTTTGCCATGGA	7129
Qy	7081	GCTGACAGTGGTGATGAGTTTTACATAGCTTCTGACACATCGGAGCCCTCTTCTCTGT	7140
Db	7130	GCTGAGCAGTGGTGATGAGTTTTACATAGCTTGTGACAACTCGGAGCCCTCTTCTCTGT	7189
Qy	7141	CTTTAGTGGAAACAGGTTTGATGATCAAGCAAAATCCTGTACACAGCCTATGGGAGATCTA	7200
Db	7190	CTTTAGTGGAAACAGGTTTGATGATCAAGCAAAATCCTGTACACAGCCTATGGGAGATCTA	7249
Qy	7201	CATGGATACCAACCCCAACTTTTCAGATCATATAGCTACCATGGTGGCTCTATCATCC	7260
Db	7250	CATGGATACCAACCCCAACTTTTCAGATCATATAGCTACCATGGTGGCTCTATCATCC	7309
Qy	7261	ACTCACCAAGCTTGTCCACATGGCCGGCGAGATTATGATGCTGGCCGGAGCGCTGGAC	7320
Db	7310	ACTCACCAAGCTTGTCCACATGGCCGGCGAGATTATGATGCTGGCCGGAGCGCTGGAC	7369
Qy	7321	TAGCCACAGACAGAGCTGTGGAAGCACTTTAGTAGCAGCAACGTGTATGCTTTTAATCT	7380
Db	7370	TAGCCACAGACAGAGCTGTGGAAGCACTTTAGTAGCAGCAACGTGTATGCTTTTAATCT	7429
Qy	7381	CTATATCTTCAAAACAAACAAACCCCATCAGCAACTCCAGAGCATCAAGTCTTCATGAC	7440
Db	7430	CTATATCTTCAAAACAAACAAACCCCATCAGCAACTCCAGAGCATCAAGTCTTCATGAC	7489
Qy	7441	AGATGTTTAAACAGCTGGCTGCTCACTTTGGATTCCAGCTACAAACGTGATCCCTGGTTA	7500
Db	7490	AGATGTTTAAACAGCTGGCTGCTCACTTTGGATTCCAGCTACAAACGTGATCCCTGGTTA	7549
Qy	7501	TCCCAACACAGACATGATGATGCCATGGAACCCCTCTACGAGCTCATCCACACAGATGAA	7560
Db	7550	TCCCAACACAGACATGATGATGCCATGGAACCCCTCTACGAGCTCATCCACACAGATGAA	7609
Qy	7561	AACGACGAGTGGGACAAACAGCAAGTCTATCTCGGGGTACAGTGTGAAGTACAGAAGCA	7620
Db	7610	AACGACGAGTGGGACAAACAGCAAGTCTATCTCGGGGTACAGTGTGAAGTACAGAAGCA	7669
Qy	7621	GCTCAAGGCCCTTTGTCACTTTAGAACCGTTTGACAGCTCTATGGTCCCAATCACCAG	7680
Db	7670	GCTCAAGGCCCTTTGTCACTTTAGAACCGTTTGACAGCTCTATGGTCCCAATCACCAG	7729
Qy	7681	CTGCCACGAGCTCCAAAGACCAAGAAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGG	7740
Db	7730	CTGCCACGAGCTCCAAAGACCAAGAAGTTTGATCCAGCGGCTCAGTCTTTGGCAAGGG	7789
Qy	7741	GGTCAAGTTTGCCCTTGAAGATGCCAGATGACACAGACATCATAGTGTGGCCAAATGA	7800
Db	7790	GGTCAAGTTTGCCCTTGAAGATGCCAGATGACACAGACATCATAGTGTGGCCAAATGA	7849
Qy	7801	GGATGGCGAGGGTGTCTGCCATCTTTGAACCATGCCCACTACTAGAGAACTCGCACTT	7860
Db	7850	GGATGGCGAGGGTGTCTGCCATCTTTGAACCATGCCCACTACTAGAGAACTCGCACTT	7909
Qy	7861	CACCATTTGATGGGTGGATACCCATTTACTTTGTGAAACCCAGGACCTTCAGAAGGTGACCT	7920
Db	7910	CACCATTTGATGGGTGGATACCCATTTACTTTGTGAAACCCAGGACCTTCAGAAGGTGACCT	7969
Qy	7921	GGCCATCTTGGGCTCTAGTGGGGGGCGGCAACCTTGAGAAATGGGGTCAACCTCACTGT	7980
Db	7970	GGCCATCTTGGGCTCTAGTGGGGGGCGGCAACCTTGAGAAATGGGGTCAACCTCACTGT	8029
Qy	7981	GTCCACAGATCAACAGTACTTAAATGGCAGACTAGACGCTACACAGACATCCAGTCCCA	8040
Db	8030	GTCCACAGATCAACAGTACTTAAATGGCAGACTAGACGCTACACAGACATCCAGTCCCA	8089
Qy	8041	GTAACGGGCACTGTCTTTGAAACACACCGTTACGGGACAAAGTGTGATGAGGAGAGGCACG	8100
Db	8090	GTAACGGGCACTGTCTTTGAAACACACCGTTACGGGACAAAGTGTGATGAGGAGAGGCACG	8149
Qy	8101	GGTCTTGAGCTGGCCGGCAGAGAGCGTTCGCGCAAGCGTGGGCCCGCAGCAGACAG	8160

Db	8150	GGTCTGTGGAGCTGGCCCGGAGAGAGCCGTGGCCCAAGCGTGGCGCCGAGCAGCAGAG	8200
Qy	8161	ACTGCGGGAAGGGGAGGAGAGCCCTGGCGGGCTGGACAGAGGGGGAGGAAGCAGAGGTGCT	8220
Db	8210	ACTGCGGGAAGGGGAGGAGAGCCCTGGCGGGCTGGACAGAGGGGGAGGAAGCAGAGGTGCT	8269
Qy	8221	GAGCAGAGGCGGGTGCAGAGGCTTACAGCGCTTTTTCGTGATCTCTGTCGAGCAGTACCC	8280
Db	8270	GAGCAGAGGCGGGTGCAGAGGCTTACAGCGCTTTTTCGTGATCTCTGTCGAGCAGTACCC	8329
Qy	8281	AGAACTGTTCAGACAGCGGCCCAACAACTCCACTTCATCAGACAGAGCGAGATGCCCGGAG	8340
Db	8330	AGAACTGTTCAGACAGCGGCCCAACAACTCCACTTCATCAGACAGAGCGAGATGCCCGGAG	8389
Qy	8341	GTGACAGAGAGGAC	8354
Db	8390	GTGACAGAGAGGAC	8403
RESULT 3			
ABN85378			
ID	ABN85378 standard; DNA; 8438 BP.		
XX	AC AC AC		
XX	ABN85378;		
XX	21-OCT-2002 (first entry)		
XX	Human NOV1, TEN-M4 like protein, coding sequence.		
DE	Human NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;		
KW	Human; NOV1; cytostatic; Cardiant; Anti-HIV; Antidiabetic; Anorectic;		
KW	Antiallergic; Haemostatic; Anti-HIV; Antidiabetic; Anorectic;		
KW	Antiasthmatic; Nephrotropic; Hepatotropic; Neuroprotective; Nootropic;		
KW	Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;		
KW	Gene therapy; NOV; cancer; heart disease; inflammation;		
KW	autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;		
KW	asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;		
KW	infection; stroke; muscular dystrophy; epilepsy; wasting disorder;		
KW	TEN-M4 like protein; chromosome 11; gene; ds.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
Key	Location/Qualifiers		
XX	4. .8395		
FT	CDS		
FT	/*tag= a		
FT	/trans except= (pos: 1138. .1147,aa:Met)		
FT	/product= "NOV1 protein"		
XX	XX		
PN	WO200255704-A2.		
XX	XX		
FD	18-JUL-2002.		
XX	XX		
PF	09-JAN-2002; 2002WO-US000554.		
XX	XX		
PR	09-JAN-2001; 2001US-0260417P.		
PR	10-JAN-2001; 2001US-0260831P.		
PR	28-FEB-2001; 2001US-0272338P.		
PR	09-MAR-2001; 2001US-0274876P.		
PR	18-APR-2001; 2001US-0284704P.		
XX	XX		
PA	(CURA-) CURAGEN CORP.		
XX	XX		
PI	Padigaru M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;		
PI	PI Zhong M, Gangolli EA, Burgess CE, Patturajan M, Vernet CAM, KA;		
PI	Taylor S, Thornev VT, Miller CE, Guo X, Soldog FL, Grosse WM;		
PI	Alsebrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;		
PI	Macdougall J, Malvankar U, Millet I, Peyman J, Smithson G;		
PI	Gunther E, Stone DJ;		
XX	XX		
DR	WPI; 2002-590674/63.		
DR	P-PSDB; ABB98401.		
XX	XX		
PT	NOVX polypeptides and encoding polynucleotides, useful for preventing or		
PT	treating NOVX-associated disorders e.g. cancer, inflammation, or		

PT Alzheimer's disease, and in chromosome mapping, tissue typing or
 XX pharmacogenomics.

Claim 9; Page 8-9; 358pp; English.

CC The present sequence is a coding sequence for a NOV protein. The NOV
 CC proteins and coding sequences are useful for treating or preventing NOV-
 CC associated disorders or in the manufacture of a medicament for treating
 CC the disorders, such as cancer, heart disease, inflammation, autoimmune
 CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
 CC IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections,
 CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
 CC and other wasting disorders associated with chronic diseases. NOV1 is a
 CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11

XX Sequence 8438 BP; 2001 A; 2434 C; 2268 G; 1735 T; 0 U; 0 Other;

Query Match	93.1%;	Score 7779.2;	DB 6;	Length 8438;
Best Local Similarity	96.4%;	Pred. No. 0;		
Matches 8111;	Conservative 0;	Mismatches 203;	Indels 100;	Gaps 11;
QY 32	GCATGGACGTGAAGGAGGAAAGCCTTACCGCTCGCTGACCGCGCGCGAGCGGAG	91		
DB 1	GCCATGGACGTGAAGGAGGAAAGCCTTACCGCTCGCTGACCGCGCGCGAGCGGAG	60		
QY 92	CGCGCTACACAGCTCTCCCGGACAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	151		
DB 61	CGCGCTACACAGCTCTCCCGGACAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	120		
QY 152	AGCTCCAGGAGACCTGAAAGGCTTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAG	211		
DB 121	AGCTCCAGGAGACCTGAAAGGCTTACGACGAGGAGGAGGAGGAGGAGGAGGAGGAG	180		
QY 212	GTCAAGGACATGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	271		
DB 181	GTCAAGGACATGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	240		
QY 272	CTGGGGAGCTGGGGCTGGAAGAGTACCGCCCTCACGGAGGAGGAGGAGGAGGAGG	331		
DB 241	CTGGGGAGCTGGGGCTGGAAGAGTACCGCCCTCACGGAGGAGGAGGAGGAGGAGG	300		
QY 332	ATTGGCTGCCCAATCGGGTACTTCCATGGGGCTGGCTCTGATGCCGACATGGAGCT	391		
DB 301	ATTGGCTGCCCAATCGGGTACTTCCATGGGGCTGGCTCTGATGCCGACATGGAGCT	360		
QY 392	GACAGGTGCTGCTCCCTGAGCACCCTGCTGGCTCTGGGGCGGAGGAGGAGGAGGAG	451		
DB 361	GACAGGTGCTGCTCCCTGAGCACCCTGCTGGCTCTGGGGCGGAGGAGGAGGAGGAG	420		
QY 452	CGCAGCTCCTGCTGTCAGCGGGGCAATTCCTCAATCTCACACTCACCGACACCGAGCAT	511		
DB 421	CGCAGCTCCTGCTGTCAGCGGGGCAATTCCTCAATCTCACACTCACCGACACCGAGCAT	480		
QY 512	GAAACACTGAGCTGATCATCCGGGGGCTGTGAGAACCAACCGCGGGCTCCCGAGCGCG	571		
DB 481	GAAACACTGAGCTGATCATCCGGGGGCTGTGAGAACCAACCGCGGGCTCCCGAGCGCG	540		
QY 572	CCGCGCGGCTCTGCGACGCCACACCCCGAAGGAGGAGGAGGAGGAGGAGGAGGAG	631		
DB 541	CCGCGCGGCTCTGCGACGCCACACCCCGAAGGAGGAGGAGGAGGAGGAGGAGGAG	600		
QY 632	CTGAACCGGGCAACTTCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	691		
DB 601	CTGAACCGGGCAACTTCAACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	660		
QY 692	CTCTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	751		
DB 661	CTCTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	720		
QY 752	CTGCTCAACAGCAACATCCCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	811		
DB 721	CTGCTCAACAGCAACATCCCTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	780		

QY 812	ACATTGAGGACAACTCTTATGATGAGGACATTTCTCGGCGCTCTCCCGCATGATGGGCT	871
DB 781	ACATTGAGGACAACTCTTATGATGAGGACATTTCTCGGCGCTCTCCCGCATGATGGGCT	840
QY 872	TACAGTGACGGGACATTTCTTTCAAGCCTGAGGACATTTCTCCCGCTCTTTTGCACACA	931
DB 841	TACAGTGACGGGACATTTCTTTCAAGCCTGAGGACATTTCTCCCGCTCTTTTGCACACA	900
QY 932	TCACAGGGTACCACTGACGTCACGACAGTGTACTCTCTCCCGCGCGAGCCCTGCGCC	991
DB 901	TCACAGGGTACCACTGACGTCACGACAGTGTACTCTCTCCCGCGCGAGCCCTGCGCC	960
QY 992	CGCAGACCTTTCGCCCGCGCGCTTTAACTCAAGAACCTTCAAGTACTGTAACTGG	1051
DB 961	CGCAGACCTTTCGCCCGCGCGCTTTAACTCAAGAACCTTCAAGTACTGTAACTGG	1020
QY 1052	AGTTCGCGAGCCTGAGCGCCATCTGTCATCTCAGCCACTCTGGTCTATCTGCTGGCATAC	1111
DB 1021	AGTTCGCGAGCCTGAGCGCCATCTGTCATCTCAGCCACTCTGGTCTATCTGCTGGCATAC	1080
QY 1112	TTTGTGCCATGCACCTGTTTGGCCTAAACTGGCACCTGACGCCGATGAGGGGAGATG	1171
DB 1081	TTTGTGCCATGCACCTGTTTGGCCTAAACTGGCACCTGACGCCGATGAGGGGAGATG	1140
QY 1172	-----TATGATCAGCGAGGACACAGCCAGTGTGGCCTGTGCTGCTGCTGCTGCTGCT	1224
DB 1141	ACGGATTTATGATCAGCGAGGACACAGCCAGCTGCGCTGTGCTGCTGCTGCTGCTGCT	1200
QY 1225	CCTATACCTCTCAGGGGCACTGGCTTAGAGACCTCTGACGAAAGGCAAGGAAACAC	1284
DB 1201	CCTATACCTCTCAGGGGCACTGGCTTAGAGACCTCTGACGAAAGGCAAGGAAACAC	1260
QY 1285	AGAAGAAAGCGGAGCTTCTTCCAGAGACAGTTCATAGATTTCTGGAGAAATTGA	1344
DB 1261	AGAAGAAAGCGGAGCTTCTTCCAGAGGCAAGTTCATAGATTTCTGGAGAAATTGA	1320
QY 1345	TGTGGGAGCGGAGCTCCAGAGATTCCTCTGCGACTTCTCGAGATCTCAAGTGT	1404
DB 1321	TGTGGGAGCGGAGCTCCAGAGATTCCTCTGCGACTTCTCGAGATCTCAAGTGT	1380
QY 1405	CATAGACCTCTCTGTCATCTGAAATTTCAATGTCTCTGGGAAAGGACGCTGTTGG	1464
DB 1381	CATAGACCTCTCTGTCATCTGAAATTTCAATGTCTCTGGGAAAGGACGCTGTTGG	1440
QY 1465	CATTATGCGAGAAAGGCTCTGCTCTTCAATACATACAGTTCATTTGTGAGCTGCT	1524
DB 1441	CATTATGCGAGAAAGGCTCTGCTCTTCAATACATACAGTTCATTTGTGAGCTGCT	1500
QY 1525	GGATGCGAGGAGCTCTTAACCCAGGAGCGGAGCTTAGAGGGAGCCCGCGCAGTC	1584
DB 1501	GGATGCGAGGAGCTCTTAACCCAGGAGCGGAGCTTAGAGGGAGCCCGCGCAGTC	1560
QY 1585	TGGGGAACTGTGCGGCTCTCAGCGCATGAGACGCTTCATCCAGTATTGATTCAGG	1644
DB 1561	TGGGGAACTGTGCGGCTCTCAGCGCATGAGACGCTTCATCCAGTATTGATTCAGG	1620
QY 1645	AATCTGCGACTTGGCTTTTCAATGACGGAAGAGTTCAGAGTTCCTTTCTCTAC	1704
DB 1621	AATCTGCGACTTGGCTTTTCAATGACGGAAGAGTTCAGAGTTCCTTTCTCTAC	1680
QY 1705	CATGCGCAATTGAGTCGGTGAATTAATGCGCCAGCAACTCTATGCAATGCTGATCAT	1764
DB 1681	CATGCGCAATTGAGTCGGTGAATTAATGCGCCAGCAACTCTATGCAATGCTGATCAT	1731
QY 1765	CTCTGGGAGCTGCGACTGCTTCTGGGTTTCTGGGCGCGAGCTGTGGGAGGCTCTCTG	1824
DB 1732	CTCTGGGAGCTGCGACTGCTTCTGGGTTTCTGGGCGCGAGCTGTGGGAGGCTCTCTG	1791
QY 1825	CCCGCTGCTCTGTAGCGGAAATGCGCAATACATGAAAGGAGATGCTGTGCGACAGTGG	1884
DB 1792	TCCAGTGTATGTAGTGGCAACGGGAGTACTTCCAAAGGCGCTGCTGCTGTTTTCAGCG	1851
QY 1885	CTGGAAGGCGCTGAGTCGATGTGCCCAACCAACAGTGTATCGATGTGGCTGCGCAAA	1944

QY 4084 CAGGGGT-----ATTACAGTGGCAAGTTTGGGCTGATCTACTTCGTGGATGGCAC 4134
 Db |||||
 QY 4072 CAGGGGTCCCCCAGGCAATACAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGCAC 4131
 Db |||||
 QY 4135 CATGATCAGACGATCGATCAGAAATGGGATCATCTCCACCTGCTCGGCTCTTAATGATCT 4194
 Db |||||
 QY 4132 CATGATCAGACGATCGATCAGAAATGGGATCATCTCCACCTGCTCGGCTCTTAATGATCT 4191
 Db |||||
 QY 4195 CACATCAGCCGGCCCACTCAGCTGTGATTTCTGTCAATGATATTTCCCAAGTAAAG- 4249
 Db |||||
 QY 4192 CACATCAGCCGGCCCACTCAGCTGTGATTTCTGTCAATGATATTTCCCAAGTAAAGCAGGT 4251
 Db |||||
 QY 4250 ----CTGGAGTGGCCCAAGACTTACCATCAACCCATGACAACTACTTTATGTCCT 4305
 Db |||||
 QY 4252 TCACCTGGAGTGGCCCAAGACTTACCATCAACCCATGACAACTACTTTATGTCCT 4311
 Db |||||
 QY 4306 CGACAAATATGTGTCTCTGCAAAATCTCTGAAACACACAGGTGGGCAATTTGCGCGGGAG 4365
 Db |||||
 QY 4312 CGACAAATATGTGTCTCTGCAAAATCTCTGAAACACACAGGTGGGCAATTTGCGCGGGAG 4371
 Db |||||
 QY 4366 GCCCATGCACTGCCAGGTCCCTGGGATGACACTTCTGCTTAAGCAAGGTGGCCATCCA 4425
 Db |||||
 QY 4372 GCCCATGCACTGCCAGGTCCCTGGGATGACACTTCTGCTTAAGCAAGGTGGCCATCCA 4431
 Db |||||
 QY 4426 CGCAACCTGGAGTCAGCCACCGCTTTGGCTGTTTCAACAATGGGTCCTGTATATTC 4485
 Db |||||
 QY 4432 CGCAACCTGGAGTCAGCCACCGCTTTGGCTGTTTCAACAATGGGTCCTGTATATTC 4491
 Db |||||
 QY 4486 TGAGACTGATGAGAAAAAGATCAACCGCATCAGGAGGTCAACCTAGTGGAGAGATCTC 4545
 Db |||||
 QY 4492 TGAGACTGATGAGAAAAAGATCAACCGCATCAGGAGGTCAACCTAGTGGAGAGATCTC 4551
 Db |||||
 QY 4546 ACTGCTTCTGGGCCCCCAGTGTGCTGATCTGATAAATGATGCCAATCTGATTTT 4605
 Db |||||
 QY 4552 ACTGCTTCTGGGCCCCCAGTGTGCTGATCTGATAAATGATGCCAATCTGATTTT 4611
 Db |||||
 QY 4606 TTCTGGAGCAGTGTTATGCAAGGATGCAAGTAAATACCCCATCTTCTTGGCTGT 4665
 Db |||||
 QY 4612 TTCTGGAGCAGTGTTATGCAAGGATGCAAGTAAATACCCCATCTTCTTGGCTGT 4671
 Db |||||
 QY 4666 GTGTGCTGATGGGAGCTCTACGTGGCGACCTTGGGAACATCCGAATTCGTTTATCCG 4725
 Db |||||
 QY 4672 GTGTGCTGATGGGAGCTCTACGTGGCGACCTTGGGAACATCCGAATTCGTTTATCCG 4731
 Db |||||
 QY 4726 CAAGAACAGCTTTCTTCAACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACCA 4785
 Db |||||
 QY 4732 GAAGAACAGCTTTCTTCAACCCAGAACATGTATGAGCTGTCTTCAACCAATTGACCA 4791
 Db |||||
 QY 4786 GGAGCTCTATCTGTTTGATACCAACCGCAAGCACTGTATACCCCAAGCTTGGCCACAGG 4845
 Db |||||
 QY 4792 GGAGCTCTATCTGTTTGATACCAACCGCAAGCACTGTATACCCCAAGCTTGGCCACAGG 4851
 Db |||||
 QY 4846 AGACTTACCTGTACAACTTCACTTACCTGGGAGCGGACATCACACTCATCACAGCAA 4905
 Db |||||
 QY 4852 AGACTTACCTGTACAACTTCACTTACCTGGGAGCGGACATCACACTCATCACAGCAA 4911
 Db |||||
 QY 4906 CAATGGCAACATGTGTAATGTCCCGCAGAGCTCTACTGGGATGCCCTCTGGCTGTGGT 4965
 Db |||||
 QY 4912 CAATGGCAACATGTGTAATGTCCCGCAGAGCTCTACTGGGATGCCCTCTGGCTGTGGT 4971
 Db |||||
 QY 4966 CCCAGATGGCCAGGTGTACTGGGTGACCATGGGACCAACATGTGCACTCAAGAGTGTGAC 5025
 Db |||||
 QY 4972 CCCAGATGGCCAGGTGTACTGGGTGACCATGGGACCAACATGTGCACTCAAGAGTGTGAC 5031
 Db |||||
 QY 5026 CACACAAAGGACAGAGTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAAC 5085
 Db |||||
 QY 5032 CACACAAAGGACAGAGTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAAC 5091
 Db |||||
 QY 5086 CAAGAAGCATGAAGACGATGGAACAACTTTATGATGACAGAGCTTTGGCCGCTGAC 5145
 Db |||||
 QY 5092 CAAGAAGCATGAAGACGATGGAACAACTTTATGATGACAGAGCTTTGGCCGCTGAC 5151
 Db |||||

QY 5146 AAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTCCTGAAGTGATACAGACAGTTTCAGT 5205
 Db |||||
 QY 5152 AAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTCCTGAAGTGATACAGACAGTTTCAGT 5211
 Db |||||
 QY 5206 GCATGTCCAGGTAGACAGCTCCAGCAAGGATGATGTCAACATAACCAACCACTGTCTGC 5265
 Db |||||
 QY 5212 GCATGTCCAGGTAGACAGCTCCAGCAAGGATGATGTCAACATAACCAACCACTGTCTGC 5271
 Db |||||
 QY 5266 CTCAGGCGCTTCTTACACACTGCTGCAAGACCAAGTCCGGAAACAGCTACTACATCGGGG 5325
 Db |||||
 QY 5272 CTCAGGCGCTTCTTACACACTGCTGCAAGACCAAGTCCGGAAACAGCTACTACATCGGGG 5331
 Db |||||
 QY 5326 CGATGCTCTTTCGGGCTGCTGCTGCAACCGCATGAGAGTGGCGCTGAGACTGAGCC 5385
 Db |||||
 QY 5332 CGATGCTCTTTCGGGCTGCTGCTGCAACCGCATGAGAGTGGCGCTGAGACTGAGCC 5391
 Db |||||
 QY 5386 CCACTTGTCTGGCTGGCAACCGTCAACCCCAACCGTGGGCAAGAGAAATGTCAACGTCGCCAT 5445
 Db |||||
 QY 5392 CCACTTGTCTGGCTGGCAACCGTCAACCCCAACCGTGGGCAAGAGAAATGTCAACGTCGCCAT 5451
 Db |||||
 QY 5446 CGACAAACGGCTCAACCTGCTGAGTGGCGCCAGGCAAGAGAGCTCGGGGCGAGT 5505
 Db |||||
 QY 5452 CGACAAACGGCTCAACCTGCTGAGTGGCGCCAGGCAAGAGAGCTCGGGGCGAGT 5511
 Db |||||
 QY 5506 CACTGTCTTTCGGGCGCGCTGCGGGT-----CACACCCGAAATCTCTATCTCT 5556
 Db |||||
 QY 5512 CACTGTCTTTCGGGCGCGCTGCGGGTCTCCAGGTTCAACCCGAAATCTCTATCTCT 5571
 Db |||||
 QY 5557 GGAATTTGATCGCGTAAACCGCACAGAGAAGATCTATGATACCCACCGCAAGTTCACCT 5616
 Db |||||
 QY 5572 GGAATTTGATCGCGTAAACCGCACAGAGAAGATCTATGATACCCACCGCAAGTTCACCT 5631
 Db |||||
 QY 5617 TCGGATTTCTAGCAGCAGCGGGGCGCCAGCTCTGCTGATCCAGCAGCAGAGGTGAA 5676
 Db |||||
 QY 5632 TCGGATTTCTAGCAGCAGCGGGGCGCCAGCTCTGCTGATCCAGCAGCAGAGGTGAA 5691
 Db |||||
 QY 5677 TGTGTCAACCTGACATCTCCCTGGGGTTACATTTGCTGATCCAGAGGGGATCAT 5736
 Db |||||
 QY 5692 TGTGTCAACCTGACATCTCCCTGGGGTTACATTTGCTGATCCAGAGGGGATCAT 5751
 Db |||||
 QY 5737 GTCTGAAGAATGGAATACGACAGCGGGGCGCCATCATCCAGATCTTCTGATGG 5796
 Db |||||
 QY 5752 GTCTGAAGAATGGAATACGACAGCGGGGCGCCATCATCCAGATCTTCTGATGG 5811
 Db |||||
 QY 5797 GAAGACATCGAGTACACATCTTAGA-----AGTCCATGCTGCTACT 5844
 Db |||||
 QY 5812 GAAGACATCGAGTACACATCTTAGAAGGAGGCTGTCAGTCCATGGTCTGCTACT 5871
 Db |||||
 QY 5845 ACACAGCCAGAGCAGTATATCTTTGAGTTGACAAAGATGACCGCTCTCTTCTGTGAC 5904
 Db |||||
 QY 5872 ACACAGCCAGAGCAGTATATCTTTGAGTTGACAAAGATGACCGCTCTCTTCTGTGAC 5931
 Db |||||
 QY 5905 GATGCCCAACGTGGCGGCGACACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAA 5964
 Db |||||
 QY 5932 GATGCCCAACGTGGCGGCGACACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAA 5991
 Db |||||
 QY 5965 CATCTATCAGCCCTTGGGCAATGCTCAGTCTACAGGACTTCACTGAGGATGGCA 6024
 Db |||||
 QY 5992 CATCTATCAGCCCTTGGGCAATGCTCAGTCTACAGGACTTCACTGAGGATGGCA 6051
 Db |||||
 QY 6025 CTTCTTCAACCTTCTAGCTGGGCACTGGCGCGAGGTGATATCAAGATATGCAAACT 6084
 Db |||||
 QY 6052 CTTCTTCAACCTTCTAGCTGGGCACTGGCGCGAGGTGATATCAAGATATGCAAACT 6111
 Db |||||
 QY 6085 GTCAAGCTGGCAGAGAGCTCTATGACACCAACCAAGGTCACTTTCACTATGACGAGAC 6144
 Db |||||
 QY 6112 GTCAAGCTGGCAGAGAGCTCTATGACACCAACCAAGGTCACTTTCACTATGACGAGAC 6171
 Db |||||
 QY 6145 GGCAGCATGTGAAGACCATCAACCTACAGAAATGAGGCTTCACTGACCATCCGCTA 6204
 Db |||||
 QY 6172 GGCAGCATGTGAAGACCATCAACCTACAGAAATGAGGCTTCACTGACCATCCGCTA 6231
 Db |||||
 QY 6205 CCGTCAGATGGGCGCCCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGGATGCT 6264
 Db |||||

Db 811 GCTTTGCCGCGAGCTGCAAAAC-----CACACCCGAGTCCGTCCAGCTGCAAGGAC 861
QY 746 AACTGGCTGCTCAACACCAATCCCTCGAGACCAAGAACTAGGCAAGCAAGCCATTC 805
Db 862 AGCTGGGTCTTGGCATATATGACCACTGGAAAGCAGGCAATTT-----905
QY 806 CTAGGGACAATTGCAAGCAACCTCAITGAGATGACATTTCTCGCGCCTCCCGCCCATGAT 865
Db 906 -----905
QY 866 GGGGCTTACAGTGACGGGCACTTCTCTTCAAGCTTGGAGCACTCCCGCTCTTCTGC 925
Db 906 -----CCTATTCAAAACAGGAACAGGTACAAGCCACTGTTTCAGT 945
QY 926 ACCACATCACAGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCTCCGCCCCGACCC 985
Db 946 ACTGCNACCCAGGATACAAATGGCATCTGGCTCTGTTTATTTCACCACTACTCGGCCA 1005
QY 986 CTGCCCCGAGCACTTTCGCCCGCCGCGCTTTAACTCAAGAACCCCTCCAGTACTGT 1045
Db 1006 CTACCTAGAAACACCTTATCAAGAAGTGCTTTTAAATTCAGAAGTCTTCAAAAGTACTGT 1065
QY 1046 AACTGGAAGTCGCGAGCCCTGAGCGCCATCGTCACTCAGCCACTCTGGTCACTCCTGCTG 1105
Db 1066 AGCTGGAAATGCACTGCACTGTGTGCGGTAGGGGTCTCGGTGCTCTGGCAATACTCCTG 1125
QY 1106 GCATCTTTTGGCCATGCACTGCTTTTGGCTTAACTGGCACTGCAAGCCGATGGAGGGG 1165
Db 1126 TCTTATTATTAGCAATGCACTCTTTTGGCTCACTTGGCGTCACTGCGAGCTAGCAAGAT 1185
QY 1166 CAGATGATGATACACGGAGGACACAGCAGAGTTGGGCTGTGCCAACCAGCACTCTCC 1225
Db 1186 GACCAATTGAGA-----ATGGAAAGTGAATTTCTGATACCATGCCAACAACACTGTG 1239
QY 1226 CTATACCCCTCAGGGGCACTGGCTTAGAGACCCCTGACAGAAAGGCAAGGAACACA 1285
Db 1240 TCATTACTCTCGAGACA-----1258
QY 1286 GAAGAAAGCCAGTAGTTCTTCCAGAGACAGTTTCAATAGATTTCTGGAGAAATGAT 1345
Db 1259 -ATGGAAATTAGTGGATTATACCAAGAAATTAACACCATAGATTCCGGAGAACTTGAT 1317
QY 1346 GTGGAGAGGCGAGCTCCAGAAATTCCTCTGGCACTTCTGGAGATCTCAAGTGTTTC 1405
Db 1318 ATTGGCCGAGAGCAATTCAGAGATTCCTCCCGGATCTTCTGGAGATCACAGCTTTC 1377
QY 1406 ATAGACCATCTGTGCACTGAAATTCATGTGTCCTGGGAAAGGCGCCCTGGTTGGC 1465
Db 1378 ATTGATCAGCCACAGTTTCTTAAATTCATATCTCTCTTCAGAAGGATGCAITGATTGA 1437
QY 1466 ATTTATGCAAGAAAGCCCTCTCTTCAATACACAGTTTGTACTTTTGGAGCTGCTG 1525
Db 1438 GTAATGCGCGGAAAGGCTTACCGCTTCCCTACTCACTGATGATCTTCTGGAGCTCCTG 1497
QY 1526 GATGCGAGAGGCTCTTAACCCAGGAGCGCGAGCCCTAGAGGGGACCCCGCCCAAGTCT 1585
Db 1498 GATGCGAGAGGCTGATTGTCAGAGAGCAGCGGAGCCTGCTTGAGACGGAGAGGCGGG 1557
QY 1586 CGGGAACTGTGCCCTCCAGCCATGACAGAGCTTCACTGATTTTGGATTCAGGA 1645
Db 1558 CGGAGGCGAGATCCGTGAGCTTCAATGAGCGCGCTTTATCCAGTACTTGGATTCGGA 1617
QY 1646 ATCTGGCACTTGGCTTTTAAATGAGGAAGAGTCAAGAGTGGTTCTCTTCTTCCACC 1705
Db 1618 ATCTGGCATCTGGCTTTTAAATGATGGGAATTAAGCAGAGAGGTTCTTTTAAATACC 1677
QY 1706 ACTGCCATGTAGTGGTGGATAAATGCCCCAGCAACTGCTATGGCAATGTGTACTGATC 1765
Db 1678 ATTGTTATAGAGTCTGTGGTGAATGTCCCCGAAATTCCTGAAATGGAGAAATGCGTT 1737
QY 1766 TCTGGCACTGCCACTGCTCTGGGTTTCTGGGCGCCGAGTCTGCGAGAGGCTCCTGC 1825

Db 1738 TCTGGAACCTTCCCATTTGTTTCCAGGATTTCTGGGTCCGATTTGTTCAAGAGCCGCTGT 1797
QY 1826 CCCGTGCTCTGTAGCGGAAATGGCCAATACATGAAGGCGAGATGCTTTGTGCCACAGTGGC 1885
Db 1798 CCAGTGTATGTAGTGGCAACGGGCACTACTTCCAAAGGCGCGTCTGCTGTGTTTCAGCGGC 1857
QY 1886 TGGAAAGCGCTGAGTGGATGTGCCCAACCAACAGTGTATCGATGTGGCTGCGCAGCAAC 1945
Db 1858 TGGAAAGGCAACCGAGTGTGATGTGCCGACTACCCAGTGTATTGACCCACAGTGTGGGGT 1917
QY 1946 CATGCACTCTCATCACGGGCACTGCACTCTGCAACCTTGCTCAACAGGCGAGAGCTGT 2005
Db 1918 CTGGGATTTGTAATCATGGGCTCTGTGCTTGCAGCTCAGGATACAAAGAGAAAGTTGT 1977
QY 2006 GAGGAAGTGGACTGATGGACCCCACTGTTTCAGGCGGGGTGTCTCGGTGAGAGGGAA 2065
Db 1978 GAAGAAGCTGACTGTATAGACCTTGGGTGTTCTAATCATGGTGTGTATGCCACGGGAA 2037
QY 2066 TGCCATTCTTTTGGGATGGGAGGACCAACTGCGAGACCCCGGCGCACATGCTTA 2125
Db 2038 TGTACTGCTAGTCCAGGATGGGAGGTAGCAATTTGTAATCTGAAGACCATGTGTCCA 2097
QY 2126 GACCAAGTGTTCAGGCCACGGAACTTCTCTCCGACACCGGGCTTTGCGAGCTGTGACCA 2185
Db 2098 GACCAAGTGTCTCGGCAACGGAACTGTTCTTCAAGAAAGTGGCTCTCTGCACTGTGACCT 2157
QY 2186 AGCTGGACTGACAGCACTGTTCTATCGAGATCTGTGCTGCCGACTGTGGTGGCCATGGC 2245
Db 2158 AACTGGAATGCGCCACAGCTGTCTCAACGAAATATGTTCTGTGGAATGTGGCTTCACAGGC 2217
QY 2246 GTGTGCTGTAGGGGGCACCTGCGCTGCGAGGATGGCTGGATGGGGGCGAGCTGTGACACAG 2305
Db 2218 GTTTCATGGGGGGGACGTGCTGTGAAGAGGCTGGACGGGCCACGCTGTAAATCAG 2277
QY 2306 CGGGCTCCACCCCGCTGTGCCGAGCATGGGACTGCGCGCAGCGCAAGTGGAGTGC 2365
Db 2278 AGAGCTCCACCCCGCTGTGCCGAGCACGGGACCTGCAAGGATGGCAAGTGTGAAATGC 2337
QY 2366 AGCCCTGCTGGAATGGGGAACACTGCACCACTCGCTCACTATCTGGAATAGGCTAGT --- 2421
Db 2338 ACCAGGCTGGAATGGAGAGCACTGCACTATCGCTCACTATTGGATAAGATAGTAAA 2397
QY 2422 -----TAAAGAGGTTGCCCTGGTTGTGCAATGGCAACGGCAGATGTACC 2467
Db 2398 GACAAAGATAGGATATAAAGAGGGTTGTCTGTGTGTGCAACAGCAATGGAAGATGTACC 2457
QY 2468 TTAGACCTGAATGTTGGCACTGGCTGCGCAGCTGGCTGGAGAGGCTGGCTGTGAC 2527
Db 2458 CTGGACCAAAATGGCGGACATTTGTGTGCCAGCTGGATGGAGAGGAGGCTGTGAC 2517
QY 2528 ACTTCCATGGAGCTGCTCGGTTGACAGCAAGACAATGATGAGATGGGCTGGTGAC 2587
Db 2518 GTAGCCATGGAGACTCTTTTGACAGATAGCAAGGACAATGAAGGGATGGACTCATTTGAC 2577
QY 2588 TGCAATGGACCTGACTGTGCTCCAGCCCTGTGCCATATCAACCCGCTGTGCTGGC 2647
Db 2578 TGCAATGGATCCCGATTGTGCTCTACAGAGTTCTGTGCCAGAAATCAGCCCTATTGTGGGGA 2637
QY 2648 TCCCCATAACCTCTGGAACATCATCCAGGAGACAGAGTCCCTGTGTGTGTCACAGCAAGACCTA 2707
Db 2638 CTGCGGGATCTCAGGACATCATTAGCCAAAGCTTCAATCGCTTCTCAGCAAGCTGCC 2697
QY 2708 CACTCCCTCTATGACCGCATCAAGTTCTCTGGGCGAGGACAGCAGCAATATCCCC 2767
Db 2698 AAATCCCTTTATGATCGAATCAGTTTCTCTTAGGATCTGATAGCCCATTTATACCT 2757
QY 2768 GGGGAGAACCCCTTTGATGGAGGCGATGTTGTGTATTCTGTGGCCAAAGTATGACATCA 2827
Db 2758 GGAGAAAGTCTCTTCAATAAGAGCCCTGCACTGTCTATCAGAGCCCAAGTACTGACTGT 2817
QY 2828 GATGAAACCCCTCGTTGGTGTGAACATCAGTTTGTCAATAACCCCTCTCTTTGGATAT 2887
Db 2818 GATGAAATCCACTATTATGGAGTAAATGTCTCGTTTTTCCATTTCCCAAGAAATATGGATAT 2877

QY 2888 ACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGCGGCAATCTCCATCATC 2947
 Db 2878 ACTATTACCCCGCAGGACGGAATGTTTACTTGGTGCAAAATGTTGGGGCCCTCTCTAACT 2937
 QY 2948 CTGCGGTTTCGAGCGGGCACTTTTCATCACACAGAGACACCCCTGTGGCTGSCCATGGAT 3007
 Db 2938 TTGGTATTGAAACGATCCCCATCTCTCACTCAGTATCATAGTGTGGATTCATGAAT 2997
 QY 3008 CGCTTCTTTGTCATGGAACCATCATCATGAGATGAGAGAGATGAGATPCCAGCTGT 3067
 Db 2998 GTCTTTTATGTGATGATACCTCTAGTATGAGAGAGAGAGAGATGAGATPCCAGCTGT 3057
 QY 3068 GACCTGAGCAATTTTGGCCGCCCAACCCAGCTGCTCTCCATCCCACTGACGTCCTTC 3127
 Db 3058 GATCTGATGATTCGTGAGGCCAAATCCCAATCATTTGTTGATCACCTTTATCCACCTTT 3117
 QY 3128 GCCAGCTCCTGTGAGAGAAAGGCCCAATTTGCGGGAATTTAGGCTTTGACAGAGGAA 3187
 Db 3118 TTCAGATCTTCTCTGAGAGAGAGTCCCATCATTCGCCGAAACACAGGTACTCCACAGGAA 3177
 QY 3188 ATCTCTATCTCTGGCTGCAAGATGAGGCTGAGTACCTGAGCAGCGGACCCCTGGGTAC 3247
 Db 3178 ACTCAATTCAGGAACAGATTTGAAACTCTCTACTTGTAGTTCCAGAGCTGAGGGTAT 3237
 QY 3248 AAATCTGCTCAGGATCAGCTCAACCCAGCCAGTCTCCCTTCAACCTCATGAAGTG 3307
 Db 3238 AAGTCAGTCTCAAGATCACCATGACCCAGTCTATTATTCATTTAATTAAGAGTT 3297
 QY 3308 CACCTCATGTTAGGCTGAGGGCGGCTCTTTCAGGAAGTGGTTGCTGAGGCCAGAC 3367
 Db 3298 CATCTTATGTTAGTGTAGTAGGAAGACTCTTCCAAAAGTGGTTTCCCTGCTCACCMAAC 3357
 QY 3368 CTGTCTCTATTATTTCAATTTGGGCAAGACAGCTCTACACAGAGGTTCTTTGGGCTT 3427
 Db 3358 TTGGCTATATCTTCAATATGGAATAAACAGATGATATATCAGAAGTCTATGGTCTA 3417
 QY 3428 TCAGAAGCTTTGTTTCGGTGGTTATGAATATGAATCTGCGCAGATCTAATCCTGTGG 3487
 Db 3418 TCTGAAGCTTGTGTAGTTGGATATGAGTATGAGTCTGTTTGGACCTGACTCTGTGG 3477
 QY 3488 GAAAGAGAACAAAGTCTCAGGCTATGAATTTGACGGTCCAGCTTGGAGATGG 3547
 Db 3478 GAAAGAGAGTGTGCTCTCAGGGCTATGAATTTGATGGATCGTCCCAATGGGTGGCTGG 3537
 QY 3548 AGCCTAGACAAACATCATGCGCTCAACATTCAAAAGTGTATCTGCAACAAAGGAATGGG 3607
 Db 3538 ACATTAGATAAATCATGCTCTGGATGTACAGAACGTTATCTGTACAGGGAACGGG 3597
 QY 3608 GAGAACAGATTTGTTCTCAGACGCTCTCTGTCTATGGAGCATCATGGCAATGGCGC 3667
 Db 3598 GAAACCCAGTTTCATCTCCAGCAGCTCCAGTCTGATGATGATCATATGGGCAATGGCGCA 3657
 QY 3668 CGGAGAGCATCTCTGCGCCAGCTGCAACGGGCTTCTGACGGCAACAGCTCTGGCC 3727
 Db 3658 AGGCGAGCATTTCTGCGCCAGTTGCAATGGTCAAGCTGATGTAACAAAGTTACTTGGCC 3717
 QY 3728 CCAGTGGCCCTCACCTGTGGCTCTGACGGGACCTCTATGTGGGTGATTTCAACTACAT 3787
 Db 3718 CCAGTGGCGTGTGTTGGGATCGATGGCAGTCTGTACGTAGTGGATTTCAACTAGCTG 3777
 QY 3788 AGAAGGATCTCCCTCTGGNAATGTCAACCAATCTCTAGAGCTGAGGATTAAGATTTC 3847
 Db 3778 CGGCGGATTTCCCTCTGGAATGTAAACAGTGTCTTAGAATCTTAAGAAATTAAGATTTT 3837
 QY 3848 AGACATAGTCAAGTCCAGACACAAATATCTACCTGCGCCACAGACCCCATGAGTGGGCC 3907
 Db 3838 AGACATAGCAGCAACCCAGCTCATAGATATCTACCTTCCACCGATCCAGTCAACGGAGAT 3897
 QY 3908 GTCTTCTTCTGACAGCAACAGCCGGGGTCTTTTAAATCAAGTCCAATGTGGTGGTG 3967
 Db 3898 CTGTACGTTTCTGACACAAACACCCGAGGAATTTATCGCCCAAAAGTCACTTACGGGGCA 3957

QY 3968 AAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGGTGACCAAGTCCCTTT 4027
 Db 3958 AAGAGCTTGTACTAAAAATGCAAGATCGTCGAGGACAGGGAGCAATGCTTCCGTTT 4017
 QY 4028 GATGACACTCGCTGCGGGATGTTGGGAAGGACACAGAGGACACACTCACCATTCCAGS 4087
 Db 4018 GACGAGCGAGATGTGGGATGAGGGAAGGCGTGGAAAGCCACACTCATGAGTCCCAA 4077
 QY 4088 GSPATTACAGTGGCAAGTTTGGGTGATCTACTTCTGTGATGGCAACCATGATCAGACGC 4147
 Db 4078 GGAATGGCAGTTGATAAGAAATGATTAATCTACTTTGTTGATGGAACCATGATTAGGAA 4137
 QY 4148 ATCGATCAGAAATGGGATCATCTCCACCTGCTGGCTCTAATGATCTCATCATCAGCCCG 4207
 Db 4138 GTTGACAAATGGAATCATATCAACTCTTCTGGGCTCTAAGATTTGACTTTCAGCCAGA 4197
 QY 4208 CCACTCAGCTGTGATTTCTGTCTGATATTTCCAGGTAAAGACTGAGTGGGCCACAGAC 4267
 Db 4198 CTTTTAACTTTGACACACAGCATGCACATCAGCCAGGTACGTCGGAATGGCCCATGAC 4257
 QY 4268 TTAGCCATCAACCCATGAGCAACTCACTTTATGTCCTCGACAAACATGTGTCTCTGAA 4327
 Db 4258 CTAGCCATTAACCTATGATTAATCTCAATTTATGTCCTGGATAAATAGTATTACAG 4317
 QY 4328 ATCTCTGAAAACCAACAGGTGCGATTGTCGGGAGGCCCATGACCTGACGCGAGTCCCT 4387
 Db 4318 ATCACTGAAAATCGTCAAGTNGCATTTGCTGTGAGCGCCCATGACCTGTCAGTTCCC 4377
 QY 4388 GGCAATGACCATCTCTGCTAAGAGGTGGCCATCCACGCAACCTTGGAGTCAGCCACC 4447
 Db 4378 GGAGTGG---AATATCTGTGGGAGACACGGGTGACAGACAACTTGAATCAGCCACT 4434
 QY 4448 GCTTTGGCTTTTACACAAATGGGTCCGTATATTTGCTGAGACTGATGAGAAAGATC 4507
 Db 4435 GCAATGCTGTCTTACAGTGGGTCTGTACATTTACTGAAACTGATGAGAAAGAAAT 4494
 QY 4508 AACCAGATCAGGAGTCACTACTAGTGGAGAGATCTCACTGCTGTGCTGGGCCCCAGT 4567
 Db 4495 AACCAGATAAGCAGGTCAACAGATGAGAGAAATCTCTTAGTGGCCGGAATACCTTCA 4554
 QY 4568 GGTGTGACTGTAAAAATGATGCCAACTGTGATTTGTTTCTGGAGACATGGTTATGCC 4627
 Db 4555 GAGTGTGACTGCAAAAATGATGCCAACTGTGATTTACTGAACTGATGAGAAAGAAAT 4614
 QY 4628 AAGGATGCAAGTTAAATACCCCATCTCTCTTGGCTGTGTGCTGATGGGAGTCTAC 4687
 Db 4615 AAGGATGCCAACTCAGTCCGCCCATCTCTCTGCTGCTTCTCCAGATGTTACATGAT 4674
 QY 4688 GTGGCCGACCTTGGGAACATCCGAATTTGGTTTATCCGGAAGAAACAAAGCTTTCTCTCAAC 4747
 Db 4675 ATTGCAGATCTAGGGAATATCCGATCCGGGCTGTGTCAAAGATAAAGCTTTTACTTAAC 4734
 QY 4748 ACCGAGACATGATGAGCTGTCTTCAACAAATGACAGGAGCTCTATCTGTTGATACC 4807
 Db 4735 TCTATGAATCTTATGAAGTTCGCTCTCCAACTGATCAAGAACTCTACATCTTTGACATC 4794
 QY 4808 ACCGCAAGCCTGTACACCCAAAGCCTGCCACAGGAGACTACCTGTACAACTTCAAC 4867
 Db 4795 AATGTTACTTCAACATATACTGTAACTTAGTCTAGTCTGTTGATTACCTTTTACATTTAGC 4854
 QY 4868 TACACTGGGAGCGGACATCATCATCATCAGACAAATGGAACAAATGTTAATGTC 4927
 Db 4855 TACACCAATGACAAATGATAATTAATCTGCTGTGACAGACAGCAATGGCAACCCCTTGAAT 4914
 QY 4928 CGCCGAGACTCTTACTGGGATGCCCTCTGCTGGTGGTGGTCCAGATGGCCAGGTGACTGG 4987
 Db 4915 AGACGGGACCAAAATCGCATGCCAGTTCGAGTGGTGTCTCTGTGATAACCAAGTATGAG 4974
 QY 4988 GTGACCATGGGACCAACAGTGCATCTCAAGAGTGTGACACAAAGGACACAGATGGCC 5047
 Db 4975 TTGACAAATAGGAACAAATGATGTTTGAAGGACATGCTGCTCAAGGACTGGAATTAGTT 5034
 QY 5048 ATGATGACATACCATGGGAATTTCCGGGCTTCTGGCAACCAAGCAATCAAAACCGATGG 5107

Db 5035 TTGTTTACTTACCATGCAATAGTGGCTTTTATGCCCACTAAAAGTGAAGAACTGGATGG 5094
 QY 5108 ACAAATTTATGATAGTACGACAGCTTTGGCGCTGACAAATGAGACCTTCCCTACTGGC 5167
 Db 5095 ACAAAGTTTTTGAATGATGACAGTGAAGTGTCTGACAAATGTTAGCTTTCCAACTGGA 5154
 QY 5168 CAGGTGACGATTTCCGAAGTATACAGACAGTTCAGTGCATGTCAGAGTAGAGACTCC 5227
 Db 5155 GTGGTCAAAACCTGCTGCGGCAATGGAAGGCTATACAGTGGACATTTAGTCACT 5214
 QY 5228 AGC---AAGGATGATGACCAATACCAACCACTGCTGCTCCCTCAGGCGCTTCTACACA 5284
 Db 5215 AGCCGAGAAGAGATGACAGATCACTTCAAACTGCTCTGATCGATTTCTTCTACACC 5274
 QY 5285 CTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGCCGATGGCTCTTTGCGGCTG 5344
 Db 5275 ATGGTTCAAGATCAGTTAAGAAACAGCTACAGATTTGGTTATGACGGCTCCCTCAGAAAT 5334
 QY 5345 CTGCTGGCCAAAGGATGGAGTGGCGCTGACAGTACAGCCCACTTGTCTGGCTGGCACC 5404
 Db 5335 ATCTACGCCAGTGGCTGGACTACACTACCAAAACAGAGCGCGTCTGGCTGGCACC 5394
 QY 5405 GTCAACCCACCGTGGGCAAGAGAAATGTACAGCTGCCATCGAACACGGCCCTCAACCTG 5464
 Db 5395 CTAACTCCGACGGTTGCCAAAGAAACATGACTTTGGCTGGCGAGAACGTCAAAACCTTG 5454
 QY 5465 GTGAGTGGCCCAAGCGCAAGAGACAGGCTGGGGCCAGGTCACTGTCTTTGGCGCGCG 5524
 Db 5455 GTGGAATGGGATTTCCGAAAGAGCAAGCCCAAGGGAAGTCAATGTCTTTGGCGCGCAAG 5514
 QY 5525 CTGGGGTGCACACCGAAATCTCTATCTCTGACCTTTGATCGGTAAACACGACAGAG 5584
 Db 5515 CTCAGGGTTAATGGCAGAAACCTCTTTTCAAGTACCTTTGATCGAACAAAGACAGAA 5574
 QY 5585 AAGATCTATGATACCAACCGCAAGTTCACTTCCGATTTCTGACGACGAGCGGGCGG 5644
 Db 5575 AAGATCTATGACGACCAACCGTAAATTTCTATGAGGATCGCTACGACAGCTGTGGGCAC 5634
 QY 5645 CCCAGCTCTGTCACCCAGCAGCAGGTGATGGTGTCAAGTGTACATCTCCCTGGG 5704
 Db 5635 CCGACTCTCTGGCTGCCAAGCAGCAAGCTGATGCCGCTCAATGTCACTTATTCACACA 5694
 QY 5705 GTTACATTTGCTGGCATCCAGAGGGCATCATGTCTCAAAAGAAATGGAATACGACAGCG 5764
 Db 5695 GGTCAATTTCCAGATCCAGCAGGACCACTAGCAGGAAATGATATGACGGACAG 5754
 QY 5765 GCGCGCATCATCCAGATCTTCTGCTGATGGGAAGACATGGAGCTACATCTTAGAG 5824
 Db 5755 GGGAGGATCGTCTCTGGCTTTTGGCTGATGGTAAACATGGAGTTACATATTTAGAA 5814
 QY 5825 AAGTCCATGGTCTGTACTACACAGCAGAGGAGTATATTTTGGATTTGCAAGAAAT 5884
 Db 5815 AAGTCCATGGTCTTCTGCTTATAGCCAGCGGCTACATCTTCGAATACGATATGTTGG 5874
 QY 5885 GACCGCTCTCTTGTGACGATGCCAAAGTGGCGGCGAGACACTAGAGACCATCCGC 5944
 Db 5875 GACCGCTGTCTGGCATCACCATGCCAGTGGCTGGCCACACCATGCGAGACCATCCGA 5934
 QY 5945 TCAGTGGCTACTACAGAAACATCTACAGCCCTCAGGGCATGCTCAGTACATACAG 6004
 Db 5935 TCCATTTGGCTACTACCGCAACATATACAAACCCCGGAAGCAACGGCTCCATCATCAG 5994
 QY 6005 GACTTCACTAGGATGGGCACCTCTTACACACTTCTACCTGGGCACTGGCGCAGGGTG 6064
 Db 5995 GACTACACGAGGAGGGCTGCTTCTACACAGCTTCTTGGGTACAACTGGAGGGTC 6054
 QY 6065 ATATACAGATATGGCAAACTGTCAAAGCTGGCAGAGCGCTCTATGACACCAAGAGTC 6124
 Db 6055 TTTATCAAATACAGAAAGCAGACTAGGCTCTCAGAAATTTTATATGATAGCACAAAGAGTC 6114
 QY 6125 AGTTTCACTATGACGACCGGCGAGGATGCTGAAGACCATCAACCTTACAGAAATGAGGC 6184

Db 6115 AGTTTACCTATGATGAAAAACAGAGGAGTCTCTAAAGACAGATAACCTCCAGAGTATGGT 6174
 QY 6185 TTCACTCCACCATCCGCTACCGTACAGATTGGGCCCCCTGATTGACCGACAGATCTTCGCG 6244
 Db 6175 TTATTTCCACCATTTAGATACAGGCAAAATGGTCCCTGATTGACAGCAGATTTCCGCG 6234
 QY 6245 TTCACTGAGGAAGGCAATGGTCAAAGCCCGTTTTCAGCTACACTATGACACAGCTTCGCG 6304
 Db 6235 TTATGTAAGATGGGATGGTAAATGCAAGATTTGACTATAGCTATGACAAACAGCTTCGA 6294
 QY 6305 GTGACCACTGCGAGGCTGTGATCAACGAGACCCCATGCCCCATTGATCTCTATCGGTAT 6364
 Db 6295 GTGACCACTGCGAGGCTGTGATCAATGAAACGCCCATGCTATGATCTGTATCAGTTT 6354
 QY 6365 GATGATGTGTCAGGCAAGACAGAGAGTGTGGGAAGTTTGGTGTCAITTTACTATGACATT 6424
 Db 6355 GATGACATTTCTGGCAAGTTTGAGCAGTTTGGGAAGTTTGGAGTTATATATATGATATT 6414
 QY 6425 AACCAAGATCATCACACAGCTGTGATGACCCACCAAGCAATTTTGTGATATGCGCAGG 6484
 Db 6415 AACCAAGATCATTTCTACAGCTGTGATGACCTATACGAACACTTTTGTGCTCATGCGCT 6474
 QY 6485 ATGAAGGAAGTCCAGTATGAGATCTTCCGCTCGCTCATGTACTGTGATGACCGTCCAGTAT 6544
 Db 6475 ATCAAGGAGATTCAAATATGAGATATTCAGGTGCTGCTCATGTACTGGAATACAAATCAGTAT 6534
 QY 6545 GATACATGGGCGGAGTAGTGAAGAGGAGCTGAAGGTAGGACCTTAGCCCAATACCACT 6604
 Db 6535 GATAACATTTGGGTCGGGTAAACCAAGAGAGATTAATAATAGGGCCCTTTGCCACACCAACC 6594
 QY 6605 CCGTACTCTATGATGATGATGCTGACGCCAGCTGCAGACAGTCTCCATCAATGACAAG 6664
 Db 6595 AAATATGCTTATGAATATGATGTTGATGACAGCTCCAAACAGTTTACCTCAATGAAAG 6654
 QY 6665 CCACTCTGGCGCTCAGCTACAGCTCAATGGGAACCTGCACTTACTGAGCCCTGGGAC 6724
 Db 6655 ATAAATGTGGCGGTCAACTACGATCTGAATGAAACCTTCCATTTACTGAAACCAAGTAAC 6714
 QY 6725 AGTGACACGGCTCACACCACTACGGTATGACATCGCGACCCGACCTCACTCGGCTGGGTAC 6784
 Db 6715 AGTGGCGCTGTGACACCCCTTCTGCTATGACCTCGAGACAGACATCACTCGACTGGTAT 6774
 QY 6785 GTGCAATCAAGATGGATGAGATGGCTTCTCTGAGGACGCGGGGGGTGATATCTTTGAG 6844
 Db 6775 GTTCAATATCGGTTGGATGAAGATGGTTTCTTACGTCAAAGGGGACGGAATCTTTGAA 6834
 QY 6845 TACAACTCAGCTGGCTCTCATCAAGGCTTACAAACCGGCTGCGAGCTGAGTGTGAGG 6904
 Db 6835 TATAGCTCAAGGGCTTCTTAACTCGAGTTTACAGTAAAGCAGTGGCTGGACAGTGTATC 6894
 QY 6905 TACCGCTACGATGGCTGGGGCGCGCTGTCCAGCAAGACAGCCACAGCCACACCTG 6964
 Db 6895 TACCGCTATGACGGCTGGGAAGGCGTGTCTTAGCAAAACCAAGCTAGGACAGACCTG 6954
 QY 6965 CAGTTCTTCTATGACAGCTGACCAACCCACCGAGTCAACCTGTACAAACCACTCC 7024
 Db 6955 CAGTTTCTTATGCTGACTTAACTTATCCCACTTAGGATTTCTATGCTACAACTATCG 7014
 QY 7025 AGCTCTGAGATCACCTCCCTCTTACTACGACTTGAAGGACACCTCTTTGCCATGGAGTGTG 7084
 Db 7015 AGTTCCAGAAATTTACCTCCCTGTATTTATGATCTCCAAGGACATCTTTTGGCATGGAAATC 7074
 QY 7085 AGCAGTGGTATGAGTGTATTAATAGCTTGTGACAACTATCGGACCCCTCTTCTGCTGTCTTT 7144
 Db 7075 AGCAGTGGGGATGAATTTCTATTTGATCGGATAACACAGGGACACCACTGGCTGTGCTC 7134
 QY 7145 AGTGAAACAGGTTTTCATGATCAAGCAAAATCTGTACACAGCCTATGGGAGATCTTACATG 7204
 Db 7135 AGTAGCAATGGGCTTATGCTGAAACAGATTCAGTACACTGATATGAGGAATCTATTTT 7194
 QY 7205 GATACCAACCCCAACTTTTCAATCATATGAGCTACCATGCTGAGCTCTATGATCCACTC 7264
 Db 7195 GACTCTAATATTGACTTTCACTGTTAATTGGATTTTCATGTTGCTGCTGTATGACCCACTC 7254

Qy	7265	ACCAAGCTTGTCCACATGGGCGCGGAGATTATGATGTCTGGCGCGAGCGCTGGACTAGC	7324
Db	7355	ACCAAAATTAACTCACTTTGGAAAGAAGATTATGACATTTTGGCAGACCGGTGGACAACA	7314
Qy	7325	CCAGCACCAAGCTGTGGAGACACCTTAGTAGCAGCAACGTCATGCTCTTTTAATCTCTAT	7384
Db	7315	CCTGACATAGAAATCTGGA--AAAGAATTGGGAAGGACCACGACTCTTTAACTTGATC	7371
Qy	7385	ATGTTTCAAAACAACACCCCATCAGCAACTCCAGGACATCAAGTGCCTCATGACAGAT	7444
Db	7372	ATGTTTATAGGAATAACAACCCCTGCAAGCAAAATTCATGCGTGAAGAATTACATCACAGAT	7431
Qy	7445	GTTAAACAGCTGGCTGCTCACCTTTTGGATTCCAGCTACACAACGCTGATCCCTGTATATCCC	7504
Db	7432	GTTAAACAGCTGGCTGGTGACATTTGGTTTCCATCTGCACAATGCTATTCTTGGATTCCCT	7491
Qy	7505	AAACCAGACATGGATGCCATGGAAACCTCTCTACAGCTCATCCACACACAGATGAAACCG	7564
Db	7492	GTTCCCAAAATTGATTTAAACAGAAACCTTCTTACGA-----ACTGTGAAGAGT	7539
Qy	7565	CAGAGTGGGACAACAGCAAGTCTPATCTCGGGGTACAGTGTGAAGTACAGAACGACGTC	7624
Db	7540	CAGCAGTGGGATGATATACGCCCATCTTCGAGTCACGACGACAGTGGCGCGCAGGCC	7599
Qy	7625	AAGCCCTTTGTCACTTAGAACGGTTTGAACGAGCTCTATAGCTCCACAATCAACGAGTGC	7684
Db	7600	AAGCCCTTCTGTCTCGCTGGGAAAGATGGCCGAGGTGCAGGTGAGCCGCGCGCGCGCGGC	7659
Qy	7685	CAGCAGGCTCCAAAGACCAAGAACTTTGATCCAGCGGCTCAGTCTTTGGCAAGGGGCTC	7744
Db	7660	---GGCGCGAGTCTCTGGCTGTGTTTCGCCACGGTCAAGTTCGTATCGGCAGAGGGCTC	7716
Qy	7745	AAGTTTGCTTGAAGGATGGCCGAGTGACCACAGACATCATCATGTGTGGCCAAATAGGAT	7804
Db	7717	ATGTGCGCGCTCAGCCAGGGCGCGTGCGAGAACCGTGTCTCAACATCGCCACACAGGAC	7776
Qy	7805	GGCGAAGGGTGTCTGCCATCTTCAACCATGCCCCACTACCTAGAGAACCTTGCATTCACC	7864
Db	7777	TGCATCAAGTGTGGCGCGTGCTCAACAACGCTTCTACCTTGGAGAACCTTGCATTCACC	7836
Qy	7865	ATTGATGGGTGGATACCCATTACTTTGTGAAACCGAGCACTTCAGAAAGGTGACCTGGCC	7924
Db	7837	ATCGAGGCAAGGACACGCACTACTTCATCAAGACCAACACGCGCGAGAGCGACCTGGC	7896
Qy	7925	ATCTGGGCTCAGTGGGGGCGCGAACCTTGGAGGATGGGTCAAGTCAAGTCACTGTGTCC	7984
Db	7897	ACGCTCGGTTTGAACAGCGGCGCGCAAGCGCTTGAGAACGGCATCAACGTGAAGGTGTG	7956
Qy	7985	CAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAGTAC	8044
Db	7957	CAGTCCACCACCGTGGTGAAACGGCAGGACGCGCAGGTTTCGCGGACGTGGAGATGCACTTC	8016
Qy	8045	GGGGCACTGTGTAACACACAGCTACGGGACACGTTGGATGAGGAGAGGACAGGGCTC	8104
Db	8017	GGCGCTGTGCGCTGCACGTGCGCTACGGCATGACCTGTGAYGAGGAGAGGCGCGCATC	8076
Qy	8105	CTGAGCTTGGCCGCGCAGAGACCGGTGCGCAACGTTGGGCCCCCGCAGCAGCAGACTG	8164
Db	8077	CTGAGCAGCGCGGACGCGCGCTCGCCGGCGCTTGGCGCGCAGCAGCAGCGCGTG	8136
Qy	8165	CGGAAAGGGAGGAAGCGCTTGGCGGCTTGACACAGGGGGGAGAAAGCAGCAGGTGCTGAGC	8224
Db	8137	CGCACGCGCAGGAGGGCGCGGCTCTGGAACGAGGGCGAGAAAGCGCAGCTGCTGAGC	8196
Qy	8225	ACAGGCGGGTCAAGGCTACGACGGCTTTTCTGTGATCTCTGTCTGAGCAGTACCCAGAA	8284
Db	8197	GCGGCAAGGTGAGGGCTACGACGGGTACTACGTACTCTCGTGGAGCAGTACCCCGAG	8256
Qy	8285	CTGTACAGCAGCGCCAAACAATCCATTATGAGACAGACCGAGATGGGCGGAGGTGA	8344
Db	8257	CTGCCACACGCGCCAAACAATCCAGTCTCTGCGGACAGCAGATCTGGCAGAGGTAA	8316

QY	8345	C	8345		
Db	8317	C	8317		
RESULT 5					
ABQ82343					
TD	ABQ82343	standard;	cdNA; 8675	BP.	
XX	AC	ABQ82343;			
XX	17-DEC-2002	(first entry)			
XX	Human	NOV15a	encoding	cdNA SEQ ID NO:35.	
XX	Human;	NOVX;	cytostatic;	neuroprotective;	anticonvulsant;
KW	cerebroprotective;	mootropic;	antidiabetic;	antiflammatory;	
KW	antiischemic;	antiarthritic;	immunosuppressive;	antiallergic;	
KW	antihaemic;	antibacterial;	protozoacide;	antihelminthic;	
KW	cancer;	leukaemia;	lymphoma;	melanoma;	neurological disorders;
KW	stroke;	ischemic cerebrovascular disease;	Alzheimer's disease;		
KW	Pick's disease;	vesicular transport disease;	cystic fibrosis;		
KW	diabetes mellitus;	Grave's disease;	gastrointestinal disorders;		
KW	ulcerative colitis;	gastric disorder;	duodenal disorder;		
KW	autoimmune disease;	allergic reaction;	autoimmune haemolysis;		
KW	rheumatoid arthritis;	gene;	chromosome 4;	ss.	
XX	OS	Homo sapiens.			
XX	Key	Location/Qualifiers			
FT	CDS	151..8328			
FT		/*tag= a			
FT		/product= "NOV15a"			
FT		/transl_except= {pos:1249..1251,aa:Ser}			
XX	WO200262999-A2.				
XX	15-AUG-2002.				
XX	31-DEC-2001;	2001WO-US049976.			
XX	29-DEC-2000;	2000US-0258928P.			
PR	02-JAN-2001;	2001US-0259415P.			
PR	04-JAN-2001;	2001US-0257885P.			
PR	20-FEB-2001;	2001US-0269814P.			
PR	09-MAR-2001;	2001US-0279863P.			
PR	29-MAR-2001;	2001US-0279832P.			
PR	29-MAR-2001;	2001US-0279833P.			
PR	13-APR-2001;	2001US-0283889P.			
PR	18-APR-2001;	2001US-0284447P.			
PR	25-APR-2001;	2001US-0286683P.			
PR	29-MAY-2001;	2001US-0294080P.			
PR	16-AUG-2001;	2001US-0312915P.			
PR	17-AUG-2001;	2001US-0313325P.			
PR	17-SEP-2001;	2001US-0322699P.			
XX	26-NOV-2001;	2001US-0333350P.			
XX	(CURA-)	CURAGEN CORP.			
XX	Spytek KA, Li L, Wolenc AP, Vernet CAM, Eisen A, Li L				
PI	Malyankar U, Shmukets RA, Tchernev VT, Spaderna SK, X				
PI	Kexuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, X				
PI	Rastelli E, Casman SJ, Boldog F, Burgess CE, Edinger				
PI	Gunther L, Smithson G, Millet I, Macdougall JR;				
XX	WPI; 2002-732706/79.				
DR	P-PSDB; ABP53586.				
XX	New NOVX polypeptides and polynucleotides useful for treat				
PT	associated disorders, such as cancers, neurological and				
PT	of vesicular transport, gastrointestinal disorders, and				
XX	diseases.				

New NOVX polypeptides and polynucleotides useful for treating NOVX-associated disorders, such as cancers, neurological disorders, disorders of vesicular transport, gastrointestinal disorders, and autoimmune diseases.

Db 1678 TCTTTTAAATACCAATGTTATAGAGTCTGTGGTGAATGTCCCGAAATTCGCATGGAAT 1737
Qy 1754 GGTGACTGCACTCTGGGACCTGCCACTGCTTCTCGGTTTCTCGGCCCGGACTGTGC 1813
Db 1738 GGAAATGCGTTCTCGAACTTGCAATGTTTTCAGAAATTCCTGGGTCGGGATTGTCA 1797
Qy 1814 AGAGCTTCTGCGGCTGCTCTGTAGCGGAATGGCCAAATACATGAAGCAGATGCTTG 1873
Db 1798 AGAGCGGCTTCTCAGTGTATGTAGTGGCAACGGGAGTACTCAAGGCGCGTGCCTG 1857
Qy 1874 TGCACAGTGTGCTGGAAGCGCTGAGTGCATGTGCCCAACCACTAGTGTATCGATGTG 1933
Db 1858 TGTTCAGCGCTGGAAGGCAACCGAGTGTGATGCGGACTACCCAGTGTATGACCCA 1917
Qy 1934 CCCTGCAGCAACATGGCACTGATCACGGGCACTGCACTGCAACCTCGGCTACAG 1993
Db 1918 CAGTGTGGGCTGTGGGATTTGATCATGGGCTCTGTGCTTGCACTCAGGATACAA 1977
Qy 1994 GGCAGAGCTGTGAGGAAGTGGACTGCATGAGACCCCACTGTCAGGCCGGGTGTCTG 2053
Db 1978 GGAGAAAGTGTGAAGAAGCTGACTGTATAGACCTCTGGGTGTTCTAATCATGTGTGT 2037
Qy 2054 GTGAGAGCGAATGCTATGTTGTGGATGGGAGGCAACAACTGCGAGACCCCAAG 2113
Db 2038 ATCCACGGGAATGTCACTGCACTCCAGATGGGAGTAGCAATTCGAAATACTGAAG 2097
Qy 2114 GCCATGCTTAGACCAAGTGTTCAGGCCACGGAACCTTCTCCGGACACCGGCTTTC 2173
Db 2098 ACCATGTGTGAGACCAAGTGTCTCGGCCACCGAAGCTATCTTCAAGAAAGTGGCTCTG 2157
Qy 2174 AGCTGTGACCAAGCTGACTGAGACAGACTGTTCTATCGAGATCTGTGCTGCCACTGT 2233
Db 2158 ACGTGTGACCTTAACCTGACTGGCCAGACTGCTCAACGAAATATGTCTGTGGACTGT 2217
Qy 2234 GGTGGCAATGCGTGTGGTATGGGGACCTGCGGCTGCGAGGATGCTGGATGGGGCA 2293
Db 2218 GGCTCACACGCGTGTTCATGGGGGAGAGTGTGCTGTGAAGAGCTGGAAGGGCCA 2277
Qy 2294 GCCTGCCACAGCGGCGCTGCCACCGCGCTGTCCGAGATGGGACTGCGCGAGCGG 2353
Db 2278 GCCTGTAATCAGAGAGCTGCCACCGCGCTGTCCGAGCAGCGGACTGCAAGGATGC 2337
Qy 2354 AAGTGCAGTGCAGCGCTGGTGGAAATGGCGAAACACTGCAACATCGCTCACTATCTGG 2413
Db 2338 AAGTGTGAATCAGCGGCTGGAATGGAGAGCACTGCATATCGCTCACTATTTGGAT 2397
Qy 2414 AGGTGAGT-----TAAAGAGGTTGCCCTGGGTTGTGCAATGGCAAC 2455
Db 2398 AAGATAGTTAAGACAAGATAGGATATAAGAGGGTTGTCTGTCTGTGCAACAGCAAT 2457
Qy 2456 GGCAGATGTACCTTAGACCTGAATGGTTGGCACTGCGCTCTGCCAGCTGGGCTGGAGAG 2515
Db 2458 GGAAGATGTACCTTGGACCAAAATGGCGGACATTTGTGTGTGCCAGCTGGATGGAGGA 2517
Qy 2516 CTGGCTGTGACATTCATCGAGAGACTGCTCGGTGACAGCAAGACAATGATGGAGAT 2575
Db 2518 CGAGCTGTGACGTAGCCATGGAGACTCTTTGCCACAGATAGCAAGGACAATGAAGGGG 2577
Qy 2576 GGCCTGTGGAGCTGATGAGACCTGACTGTGCTGCTCCAGCCCTGTGCCATATCAACCC 2635
Db 2578 GGACTCATTTGACTGATGATTCGGATTTGCTGTGCTTACAGAGTTCTTGCAGAAATCAGCC 2637
Qy 2636 CTGTGCTTGGCTGCCCTTAAACCTTGTGAACATCATTCAGGAGACACAGGTCCCTGTGTCA 2695
Db 2638 TATTGTGCGGAGCTGCCGATCTCAGGACATCATTAGCCAAAGCCTTCAATCGCCTTCT 2697
Qy 2696 CAGCAGAACCTTACACTCTCTATGACCGCATCAAGTTCCTCGTGGGCGGAGCAGCAG 2755
Db 2698 CAGCAAGCTGCAAAATCTTTATGATCGAATCAGTTTCTTTATAGATCTGATAGCACC 2757
Qy 2756 CACATAATCCCGGGAGAACCCCTTTGATGAGGAGCATGTTGTGTATTCTGTGGCCAA 2815
Db 2758 CATGTTATACCTGGAGAAAGTCTTTCAATAAGAGCCTTGATCTGTCTCATCAGAGGCCAA 2817

Qy 2816 GTGATGACATCAGATGGAACCCCTGTTGTGTGAACATCAGTTTGTCTAATACCCCT 2875
Db 2818 GTACTGACTGCTGATGGAACCTCACTTATGTGAGTAAATGTCTGTTTTTCAATTACCCA 2877
Qy 2876 CTCCTTTGGATATACAAATCAGCAGGCAAGATGACGCTTTGACTTGGTGACAAATGGCGC 2935
Db 2878 GAAATATGAAATATCTATTTACCCGCAAGACGAAATGTTTGACTTGGTGCAAAATGGTGG 2937
Qy 2936 ATCTCCATCATCTCGGTTGAGCGGCACTTTTCATACACAGGAGCACCCCTGTGG 2995
Db 2938 GGCCTCTCTAACTTGGTATTTGAAGATCCCATTCCTCACTCAGTATCATACTGTGTGG 2997
Qy 2996 CTGCCATGGATGCTTCTTTGCTATGGAACCATCATCATCAGATGAGAGGAGATGAG 3055
Db 2998 ATTCCATGGAATGTCTTTTATGTGATGATACCCCTAGTATGGAAGAAAGAGATGAC 3057
Qy 3056 ATTCCAGCTGTGACCTGAGCAATTTTGCCTCCGCCCAACCCAGTCTCTCTCCATCCCA 3115
Db 3058 ATCCAGCTGTGATCTGAGTGAATTCGTGAGGCCAAATCCCATCATTTGTGTATCACCT 3117
Qy 3116 CTGACGTCTTCCGAGCTCTCTGTGAGAGAAAGCCCATTTGTGCCGGAATTCAGGT 3175
Db 3118 TTATCCACCTTTTTCAGATCTTCTCTGAAGACAGTCCCATCATTTCCCGAAACACAGT 3177
Qy 3176 TTGACGAGGAAATCTCTATCTCTGGCTGCAAGATGAGCTGAGCTACCTGAGCAGCGG 3235
Db 3178 CTCCAGAGAAACTACATTCAGGACAGATTTGAAACTCTCTACTTGTGATTTCCAGA 3237
Qy 3236 ACCCTTGGTACAATCTCTCTGAGGATCAGCTCACCCACCCGACCATCCCTTTCAAC 3295
Db 3238 GCTGACAGGTATAAGTCACTTCTCAAGATCACATGACCCAGTCTATTATTCCATTAT 3297
Qy 3296 CTCATGAGGTGACCTCATGTTAGGTGAGGCGCGCTCTTCAGGAAGTGTTCGCT 3355
Db 3298 TTAATGAAGTTTCATCTTATGTTAGTGTGATGAGAGACTCTTCCAAAGTGTTCCT 3357
Qy 3356 GCAGCCCCAGACTGCTCTATTATTTCATTTGGGCAAGACAGAGCTCTACACCAAG 3415
Db 3358 GCTCACCAAACTTGGCTTACTTTTCATATGGGATAAAACAGATGCATATAATCAGAAA 3417
Qy 3416 GTGTTTGGCTTTCAGAACCTTTGTTCCGTGGGTATGATATGAATCTTCCCTGCCAGAT 3475
Db 3418 GTCTATGTGTTCTATGAACTGTTGTGTCAGTTGATATGATGATGAGTGTGTTGGAC 3477
Qy 3476 CTAACTCTGGGAAAAAAGAACAAAGTCTGCAAGGCTATCAAAATTCACGCGTCCAAG 3535
Db 3478 CTGACTCTCTGGGAAAAAGAGACTGCCATTTCTGCGAGGCTATGAATTTGATGGTCCAAC 3537
Qy 3536 CTGAGGATGAGCCTAGACAAACATCATGCGCTCAACATTCAAAGTGGTATCTGTGAC 3595
Db 3538 ATGGTGGCTGGACATTAGATAAATCAATCACTGCTGGATGTACAGAACGTTATCTGTAC 3597
Qy 3596 AAGGGAATGGGGAACCAAGTTTGTCTCAGAGCCTCTCTCATTTGGAGCATCATG 3655
Db 3598 AAGGGAACCGGGAACCAAGTTTCTCTCCAGAGCTCTCCAGTCTGTGATGATCATG 3657
Qy 3656 GGCATTTGGGCGGGAAGAGCATCTCTGCCCCAGCTGCAACCGGCTTCTGTCAGGCAAC 3715
Db 3658 GGCATTTGGGCGAAGGCGCAGCATTTCTGCCCCAGTTGCAATGGTCAAGCTGATGTAAC 3717
Qy 3716 AAGCTCTGGCGCCAGTGGCCCTCACCTGTGCTGTGAGCGGAGCTCTATGTGGTGAT 3775
Db 3718 AAGTTACTGGCCCAAGTGGCGCTAGCTTGTGGGATGATGAGGAGTCTGTACGTAGCGAT 3777
Qy 3776 TTCACCTACTATGAGAGGATCTTCCCTCTGGAATGTCAACAAATCTTAGAGTGGAG 3835
Db 3778 TTCACCTAGTGGCGGATATTCCTTCTGGAATGTAAACAGTGTCTTAGAATTAAGA 3837
Qy 3836 AATAAAGATTTCAGACATAGTCCAGTCCAGCACAAATCTACTGCGCCACAGACCC 3895
Db 3838 AATAAAGATTTAGACATAGCAGCAACCCAGCTCATAGATACTTACTTCTCAACGGATCCA 3897

QY 3896 ATGAGTGGGCGCTTCTCTTCTGACAGCAACAGCGCGGGTCTTTTAAATCAAGTCC 3955
 Db 3898 GTACGGGAGATCTGTACGTTTCTGACACAAACCCCGCAGAAATTTATCGCCCAAGTCA 3957
 QY 3956 ACTGTGTGTGAAGACCTTGTCAAGAACTCTGAGTGTGTGGGGGACAGGTGACCAAG 4015
 Db 3958 CTTACGGGGGCAAAAGACTTGAACATAAATGACAAAGTCTGTCGAGGAGACAGGGGACAA 4017
 QY 4016 TGCCTCCCTTTGTAGTACACTCGCTGCGGGGATGCTGGGAAGGCCACAGAGGCACACTC 4075
 Db 4018 TGCCTCCCTTTGACAGGCGAGATGTTGGGATGGAGGAAAGCGCTGGGAAGCCACTC 4077
 QY 4076 ACCAATCCAGGGGTATTACAGTGGACAAAGTTTGGGTGATCTACTTCTGGTGGATGGCACC 4135
 Db 4078 ATGATGCCAAAGGAATGGCAGTTGATGAAGAAATGGATTAATCTACTTTGTGTGGAACCC 4137
 QY 4136 ATGATCAGACGATCAGATGGAATGGATCATCTCCACCTGCTCGGCTCTAATGATCTC 4195
 Db 4138 ATGATTAGGAAGTTGACCAAAATGGAAATCATATCAACTCTTCTGGGCTCTAAGATTG 4197
 QY 4196 ACATCAGCCCGGCCACTCAGCTGTGATTTCTGTCATGATATTTCCAGGTAAGACTGGAG 4255
 Db 4198 ACTTCAGCCAGACTTTAACTTTGTGACACAGCATGCACATCAGCCAGGTACGTTCTGGAA 4257
 QY 4256 TGGCCACAGACTTAGCCATCAACCCATGAGCAACTCAGTTTATGTCTCGACAAAT 4315
 Db 4258 TGGCCCACTGACCTAGCCATTAACCTATGATTAATCCATTTATGTCTGGATTAAT 4317
 QY 4316 GTGCTCTGCAAACTCTGAAACCAACAGAGTGGCATTTGCGCGGGAGGCCCATGAC 4375
 Db 4318 GTAGTTTACAGATCACTGAAATCGTCAAGTTTCGATTTGCTGTGACGCGCCCATGAC 4377
 QY 4376 TGCAGGTCCCTGGCATTTGACCACTTCTGCTGAAGAGGTGGCCATCCAGCAACCTG 4435
 Db 4378 TGTCAAGTTCCGGAGTGG---AATATCCTGTGGGAAGCACGCGGTGCAGACAACACTG 4434
 QY 4436 GAGTCAGCCCGCTTGGCTGTTTACAGAAATGGGGTCTGTATATTGTGTGAGACTGAT 4495
 Db 4438 GAATCAGCCACTGCAATTTGCTGTCTCAAGTGGGGTCTGTACATTTACTGAACTGAT 4494
 QY 4496 GAGAAAAGATCAACCGCATCAGCAGGTCAACACTAGTGGAGAGATCTCACTCGTTGCT 4555
 Db 4498 GAGAAGAAATTAACCGGATAAGCAGGTCAACACAGATGGAGAAATCTCTTAGTGGCC 4554
 QY 4556 GGGCCCCAGTGGCTGTGACTGTAAATGATGCCAACTGTGATTTGTTTCTGAGAC 4615
 Db 4558 GGAATACCTTCAGAGTGTGACTGTCAAAATGATGCCAACTGTGACTGTTTACCAGAGTGA 4614
 QY 4616 GATGTTATGCCAAGGATGCAAGATTAATACCCCATCTTCTTGGCTGTGTGTGTGAT 4675
 Db 4618 GATGCTACGCCAAGGATGCCAACTCAGTCCGCCCATCTCTCTGGCTGTCTCCAGAT 4674
 QY 4676 GGGGAGCTCTACGTGGCGGACTTGGGAAATCCGAAATTCGGTTTATCCGGAAGACAG 4735
 Db 4678 GGTACACTGTATATTGAGATCTAGGGAATATCCGATCCGGGCTGTGTCAAAAGAAATAG 4734
 QY 4736 CTTTCTCTCAACCCAGACATGTATGAGTGTCTTCAACAAATGACAGGAGCTCTAT 4795
 Db 4738 CTTTACTTAATCTATGAACTCTATGAAAGTTGCTCTCAACTGATCAAGAACTCTAC 4794
 QY 4796 CTGTTTATACCAACCGGACAGCACTGTACACCAAAAGCCTGCCACAGAGACTACCTG 4855
 Db 4798 ATCTTTGACATCAATGGTACTCAACCAATATCTGTAAGTTTGTACTGTGTGATTACCTT 4854
 QY 4856 TACAACCTCACTACACTGGGACGGGACATCACTCATCAGACACAAATGGCAAC 4915
 Db 4858 TACAATTTTACTACGAATGAAATGATATTACTGCTGTGACAGACAAATGGCAAC 4914
 QY 4916 ATGTTAATGTCCCGGAGACTTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGC 4975
 Db 4918 ACCCTTAGAATTAGACGGGACCAAAATCGCATGCCAGTTCCGAGTGGTGTCTCTGATAAC 4974
 QY 4976 CAGGTGTACTGGGTGACCATGGGACCAACAGTGCATCAAGAGTGTGACCACACAGGA 5035

Db 4975 CAAGTGATATGTTGACAATAGGAACAAATGGATGTTTGAAGGCGATGACTGCTCAAGGA 5034
 QY 5036 CACGAGTTGGCCATGATGACATACCATGCAATTCGGGCTTCTGGCAACCAAAAGCAAT 5095
 Db 5038 CTGGAATTAGTTTGTCTTACTTACCATGGCAATAGTGGCCTTTTAGCCACTTAAAGTGTAT 5094
 QY 5096 GAAAACGGATGGACAACATTTTATGATGACGACAGCTTTGGCCGCTGACAAATGTGACC 5155
 Db 5098 GAACTGTGATGACAACCGTTTGTGACTATGACAGTGAAGTCTGCTGACAAATGTATAG 5154
 QY 5156 TTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGTATCAGACAGTTTCACTGATGTCAG 5215
 Db 5158 TTCCAACTGGAGTGTGACAAACCTGATGGGACATGGAACAGGCTTATCACAGTGGAC 5214
 QY 5216 GTAGAGACTCCAGC---AAGGATGATGTCAACATAACCAACCACTGCTGTGCTCAGGC 5272
 Db 5218 ATTGAGTCACTAGCCGAGAGAAGATGTGAGATCACTTCAATCTGTCTCTCGATCGAT 5274
 QY 5273 GCCTTCTACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACTATCGGGGCGGATGGC 5332
 Db 5275 TCTTTCTACCATGTTTCAAGATCAGTTAAGAAAACAGCTACCAAGATTTGGTTATCAOGGC 5334
 QY 5333 TCCTTTGGGCTGCTGCTGGCCAAAGGATGGAGTGGGCTGCGAGCTGACAGCTGACGCCACTG 5392
 Db 5335 TCCCTCAGAAATTTACTACGCAAGTGGCTGGACTCACTACCAACAGAGCCGACGTT 5394
 QY 5393 CTGGTGGCACCGTCAACCCCAAGGCAAGGAAATGTCAAGCTGCTGCTGCTGCTGCTGCTG 5452
 Db 5395 CTGGTGGCACCGTCAACCCCAAGGCTTGCACAAAGAAACATGACTTTTGGCTGGCGAGAC 5454
 QY 5453 GGCCTCAACCTGTGGAGTGGCCGACGCAAGAGAGGCTCGGGCCGAGGTCACTGTCTG 5512
 Db 5455 GGTCAAAACTTGGTGGAAATGGAGATTCGAAAAGAGCAAGCCCAAGGAAAGTCAATGTCT 5514
 QY 5513 TTTGGGCGCGGCTGCGGGTGCACAAACCGAAATCTCTCTGCTGCTGCTGCTGCTGCTGCTG 5572
 Db 5515 TTTGGGCGCGGCTGCGGGTAAATGGCGAGAAATCTCTCTTCTGCTGCTGCTGCTGCTGCTG 5574
 QY 5573 ACACGCAACAGAGAAGATCTATGATGACCAACCGAAGTTTCACTTGGGATTTCTGTAGAC 5632
 Db 5575 ACAAAGACAGAAAATCTATGACACCAACCGTAAATTTCTACTGAGGATCCGCTAGCAC 5634
 QY 5633 CAGGCGGGCGGCGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5692
 Db 5635 AGCTGCGGACCCGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5694
 QY 5693 TACTCCCTGGGGTTACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5752
 Db 5695 TATTCACTCAGAGTCAAAATTTGCGAGCATCCAGCGAGCACCACTAGCGAGAAATAGAT 5754
 QY 5753 TACGACAGGGCGGCGGCGCATCACATCCAGGATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 5812
 Db 5755 TATGACGACAGGGGAGGATCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5814
 QY 5813 ACATATTTAGAGAAGTCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5872
 Db 5815 ACATATTTAGAAAAGTCCATGCTTCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCTGCTT 5874
 QY 5873 TTGCAAGAAATGACCGGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5932
 Db 5875 TACGATATGGGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5934
 QY 5933 GAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCTGAGGGCAATGCCC 5992
 Db 5935 CAGACCATCCGATCCATTTGGGCTACTACCGCAACATATACAAACCCCGGAGAGCAAGCCC 5994
 QY 5993 TCAGTATACAGGACTTCACTGAGGATGGGCACTCTCTTCTTCACTTCTTCTTCTTCTTCTTCTTCT 6052
 Db 5995 TCCATCATCAGGACTACACAGGAAAGGCTGCTTCTTCAAAAACAGCTTTCTTCTGGGTACA 6054
 QY 6053 GCGCGCAGGGTGATATCAAGTATGGCAAACTGTCAAAAGCTGGGAGAGACGCTCTATGAC 6112

Db 6055 AGTCGGAGGGTCTTTATTAACAATACAGAAAGCAGACTAGGCTCTCAGAAAATTTTATATGAT 6114
 Qy 6113 ACCACCAAGGTCAGTTTCACTATGACGAGACGCGCATGCTGAGAGACCATCAACCTA 6172
 Db 6115 AGCACAAGATCAGTTTACTATGATGAACAGCAGAGTCTTAAGACAGATAAACCTC 6174
 Qy 6173 CAGAAATGAGGGCTTCACTGACCAATCCGCTACCGTCAGATGCGGCCCTGATGACCGA 6232
 Db 6175 CAGAGTGATGGTTTATTTGACCACTAGATACAGGCAAAATGGTCCCTGATGACAGG 6234
 Qy 6233 CAGATCTTCGGCTTCACTGAGGAGGATGCTGACAAAGCCGTTTGGACTACACTATGAC 6292
 Db 6235 CAGATTTTCGCTTTAGTGAAGATGGATGATGAATGCAAGATTTGACTATAGCTATGAC 6294
 Qy 6293 AACAGCTTCGGGTGACAGCATGACAGCTGTGATCAACGAGACCCCACTGCCATTTGAT 6352
 Db 6295 AACAGCTTCGAGTGACAGCATGACGGGTGTGATCAATGAACGCGCACTGCCATTTGAT 6354
 Qy 6353 CTCTATCGCTATGATGCTGACGCAAGACAGACAGAGTTCGGAGTTTGGTGTCAATT 6412
 Db 6355 CTGATCAGTTTGAAGACATTTCTGGCAAGTTGAGCAGTTTGGAAAGTTTGGAGTTATA 6414
 Qy 6413 TACTATGACATTAACACAGATCATCAACACAGCTGTCTATGACCCACCAAGCAATTTTAT 6472
 Db 6415 TATTATGATATTAACAGATCAATTTCTACAGCTGTANTGACCTATACGAGACCTTTGAT 6474
 Qy 6473 GCATATGGCAGGATGAAGAGTGCAGTATGAGATCTTCGCTGCTCATGTAAGTATG 6532
 Db 6475 GCTCATGGCCGTATCAAGAGGATTTCAATATGAGATATTCAGTGGCTCATGTACTG 6534
 Qy 6533 ACCGTCAGATGATATAATGCGGCGAGTAGTGAAGAGGAGCTGAGAGTGAAGACCTTAC 6592
 Db 6535 ACAATTCAGTATGATATCATGGTTCGGGTACACGAGAGAGATTAATATAGGGCCCTTT 6594
 Qy 6593 GCCAATACCACTCGCTACTCTATGAGTATGATGCTGACGCCAGCTGACAGACAGTCTCC 6652
 Db 6595 GCCAACAACCAAAATATGCTTATGAATATGATTTGATGACAGCTCCAAACAGTTTAC 6654
 Qy 6653 ATCAATGACAGCCACTTCGCGCTACAGTACGACCTCAATGGACCTGCACTACTG 6712
 Db 6655 CTCATGAAGAATAATGTCGGGTCAACTACGATCTGAATGAAACCTCCCAATTTACTG 6714
 Qy 6713 AGCCCTGGGAACAGTGCAGCGCTCACACCACTACGCTATGACATCCCGACCGCATCACT 6772
 Db 6715 AACCCAAAGTAAACAGTGCAGCTGACACCCCTTCGCTATGACCTGCGAGACAGATCACT 6774
 Qy 6773 CGGCTGGGTGACGTGCAATCAAGATGATGAGATGCTTCTTGAGGACAGCGGGCGGT 6832
 Db 6775 CGACTGGGTGATGTTCAATATCGGTGGATGAAGATGTTTCTTACGTCAAAGGGGCAAG 6834
 Qy 6833 GATATCTTTGATGATCAACTCAGCTGGCTGCTCATCAAGGCTTACAAAGGCTGGCAGC 6892
 Db 6835 GAAATCTTTGATATAGCTCAAGGGCTTCTAATCGAGTTTACAGTAAGGAGCTGGC 6894
 Qy 6893 TGGAGTGTACGTTACCGTACGATGCGCGCGCGGTGTCAGACCAAGAGCAGCCAC 6952
 Db 6895 TGGACAGTATCTACCGTTATGACGGCTTGGAAAGGCTGTTCTAGCAAAACAGTCTA 6954
 Qy 6953 AGCCACCACTGCAAGTCTTCTATGACAGCTGACCAACCCCAAGGTCACCCACCTG 7012
 Db 6955 GGAAGACACCTGCAAGTCTTCTATGCTGACTTAATCTTATCCCACTAGGATTAATGCTC 7014
 Qy 7013 TACAACCACTCGAGTGTAGATCACCTCCCTCTACTACGACTTGAAGGACACCTCTTT 7072
 Db 7015 TACAACCACTCGAGTTCAGAAATTAACCTCCCTGTATATGATCTCCAAGGACATCTTTT 7074
 Qy 7073 GCCATGAGAGTGCAGTGTGATGAGTTTATACAGCTTGTGACAAATCATCGGACCCCT 7132
 Db 7075 GCCATGGAATCAGCAGTGGGGATGAATCTTATATTCATCGGATATACAGAGGACACCA 7134
 Qy 7133 CTTCCTGTCTTGTGAAACAGTGTGATGATCAAGCAATCTCTGTACAGCCTATGGG 7192
 Db 7135 CTGGCTGTCTCAGTAGCAATGGCTTATGCTTGAACAGATTCAGTACACTGATATGGG 7194

Qy 7193 GAGATCTATGATACCAACCCCAACTTTTTCAGATCATCATAGGCTAACCATGGTGCCTC 7252
 Db 7195 GAAATCTATTTGACTCTAAATATTGACTTTCAACTGGTAAATTTGATTTTCAATGGCTG 7254
 Qy 7253 TATGATCAGTACCAAGAGTGTCCATGATGGCGCGAGATTAATGATGCTGCTGCCGGA 7312
 Db 7255 TATGACCCACTCACCAAAATTAATCCACTTTGGAGAAAGAGATTAATGATTTTGGCAGA 7314
 Qy 7313 CCGTGGACTAGCCAGACACAGAGCTGTGGAAGCACTTAGTAGCAGCAACGCTCATGCT 7372
 Db 7315 CGTGGACAAACCTGACATAGAAATCTGGA---AAGAAATTTGGAAGGCCAGCTCCT 7371
 Qy 7373 TTTAATCTCTATATGTTTCAAAAACCAACCCCATCAGCAACTCCCAAGACATCAAGTGC 7432
 Db 7372 TTTAATCTCTATGTTTGTAGAAATTAACCAACCCCTGCAAGCAAAATCCATGACGTGAAGAT 7431
 Qy 7433 TTTCAATGAGAGTTTAAACAGCTGGCTCAGCTTTTGGATTCCAGCTACACAACTGATC 7492
 Db 7432 TATATCAGAGATTTAAACAGCTGGCTGGTGAATTTGGTTTCCATCTGCACATGCTATT 7491
 Qy 7493 CTTGTTATCCCAACCAAGACATGATGCCATGGAACCTCTCTACGAGCTCATCCACACA 7552
 Db 7492 CTTGATTTCCCTGTTTCCCAAAATTTGATTTAAACAGAACTTCTTACGA-----A 7539
 Qy 7553 CAGATGAACACGACAGAGTGGACAAACAGCAAGTCTATCTCTCGGGGTACAGTGAAGTA 7612
 Db 7540 CTTGTGAAGAGTCAGCAGTGGGATGATATACCGCCCATCTTCGGAGTCCAGCAGCAAGTG 7599
 Qy 7613 CAGAAGCAGCTCAAGGCTTTGTACCTTAGAACGGTTTGACCAAGCTTATGCTGCCACA 7672
 Db 7600 GCGCGGAGCCAAAGGCTTCTCTGTCGCTGGGAAGATGGCGAGTGCAGTGAAGCCG 7659
 Qy 7673 ATCAGCAGTGCACAGAGCTCCAAAGACCAAGAGTTTGCATCCAGCGGCTCAGTCTTT 7732
 Db 7660 CGCCGGGCGCGC---GGCGCGCAGTCTCGCTGTGTTTCCGACACGGTCAAAGTCCGTGATC 7716
 Qy 7733 GCAAGAGGGGTCAAGTTTTCCTTGAAGATGCGCGAGTGACACAGACATCATCAGTGTG 7792
 Db 7717 GCAAGAGGGTCAATGCTGCGCTGAGCGGCGGTGACAGCCGCTGACACCAAGTCTCAATC 7776
 Qy 7793 GCAATGAGGATGCGGAGGTTGCTGCCATTTTGAACCAATGCGCCCATCTAGAGAAC 7852
 Db 7777 GCAACAGAGGACTGCATCAAGGTGGCGCGTCTCAACACCGCTTCTTACCTGGAGAAC 7836
 Qy 7853 CTGCACTTCAACATTTGATGGGTGATACCCATTTACTTTGTGAACACAGGACCTTCAGAA 7912
 Db 7837 CTGCACTTCAACATCGAGGGCAAGGACACGCACTACTTCAATCAGACCAACCGCCCGAG 7896
 Qy 7913 GGTGACCTGGCCATCTGCGGCTCAGTGGGGGCGCGAAACCTCGAGAAATGGGGTCAAC 7972
 Db 7897 AGCGACCTGGGACGCTCGGTTGACAGCGGCGCAAGGCGCTGGAGAACGSCATCAAC 7956
 Qy 7973 GTCACTGTGTCAGATCAACAGTACTTAAATGGCAGGACTAGAGCTACACAGACATC 8032
 Db 7957 GTGACGGTGTGCTCAGTCCACACCGTGTGTGAACCGGACGACGCGAGGTTCCGCGAGCTG 8016
 Qy 8033 CAGCTCCAGTACGGGCACTGTGTTGAACACACGCTACGGGACAACTTTGGATGAGGAG 8092
 Db 8017 GAGATGAGTTCGGCGCTGCGCTGACGCTACGCGCTACGCGCTGACCGGAGGAG 8076
 Qy 8093 AAGGCAAGGCTCCTGAGCTGGCGCGGAGAGCGCTGCGCCAAAGCGTGGCGCCCGAG 8152
 Db 8077 AAGGCGCGCATCTGGAGCAGCGCGGAGCGCGGCTCGCCGCGGCTGGCGCGCGAG 8136
 Qy 8153 CAGCAGAGACTCGGGAGGGGAGGAGGCTCGGGCTCGGAGAGGGGAGGAGAGAG 8212
 Db 8137 CAGCAGCGGCTCGGAGCGCGGAGGAGGCGCGGCTCTGGACGAGGGGAGGAGCGG 8196
 Qy 8213 CAGGTGCTGAGCAGAGGCGGCTGCAAGGCTACGACGCGTCTTCTGATCTCTCTCGAG 8272
 Db 8197 CAGCTGCTGAGCGCGGCAAGGTCAGGGCTACGACGCGTACTAGTACTCTCTCGGTGAG 8256

Db 1217 ACAGTTTCTTAAATCAATATCTCTCTCAGAGGATGCATTGATTGGAGTATATGGCCG 1276
Qy 1477 AAAAGGCTCCCTCCTTCAACATACAGTTTGAATTTGTGAGAGTGTGATGAGCAGAG 1536
Db 1277 GAAAGGCTTACCGCTTCCCATCTAGTATGATTTGCTGAGCTCTGATGAGCAGAG 1336
Qy 1537 GCTCTAACCCAGAGAGCGGAGCTTAGAGGGACCCGCGCAGTCTCGGGAACTGT 1596
Db 1337 GCTGATTGCCAGAGAGCAGCGAGCTTGTGAGACGAGAGAGCCGGGCGGAGGGAG 1396
Qy 1597 GCCCCCTCCAGCATGAGCAGAGCTTCAATCAGTATTTGATTCAGGAATCTGGCACTT 1656
Db 1397 ATCCGTGAGCCTTCATGAGCGCGCTTATCCAGTACTTGGATTTCTGGAATCTGGCATCT 1456
Qy 1657 GCTTTTATACATGACGGAAGAGTCAAGAGTGGTTTCTTCTCACCCTGCCATTGA 1716
Db 1457 GCTTTTATATGATGGAAAAATGAGAGAGTGTCTTTTAATACCATTTGTTATAGA 1516
Qy 1717 GTCGGTGGATAACTGCCCGAGCAACTGCTATGGCAATGGTGAATGCTCTGGGACCTG 1776
Db 1517 GTCTGTGTGAATGTCCCGAAATGGCAATGGAGATGCGTCTTCTGGAATCTG 1576
Qy 1777 CCATGCTTCTGGTTCCTGGGCGGAGTGTGGCAGAGCTCTCTGCCCCGCTCTG 1836
Db 1577 CCATGTTTTCAGAGTTCTGGGTCGGATTTGTCAGAGCGCGCTGTCCAGTGTATG 1636
Qy 1837 TAGCGAAATGGCCAAATACATGAAGGAGATGCTTGTGCCAGTGGCTGGAAGGCGC 1896
Db 1637 TAGTGGCAACGGCAGTACTCCAGGCGCTGCTGTGTTTCAAGCGCTGGAAGGCGAC 1696
Qy 1897 TGAGTGGATGTGCCCAACCAAGTGTATCGATGTGGCTGTGAGCAACCAATGGCACTG 1956
Db 1697 CGAGTGTATGTGCCGATACCCAGTGTATTGAACCAAGTGTGGGGTGTGGGATTTG 1756
Qy 1957 CATCAGGGCACTGCATCTCCAACTGCTGACAGGCGAGAGCTGTGAGGAAGTGA 2016
Db 1757 TATCATGGCTCTTGTGCTTCCAACTCAGGATACAAAGGAAAAAGTTGTGAAGAGCTGA 1816
Qy 2017 CTGATGACGCCCAATGTTTCAGGCGGGGTGTCTGCGTGTGAGAGCGAAATGCCATTCCT 2076
Db 1817 CTGTATAGACCTCGGTGTTCTAATCATGTGTGTGTATCCAGGGGAATGTCACTGAG 1876
Qy 2077 TGTGGATGGAGGACCACTGAGAGACCCAGGCGCACATGCTTGAACAGTCTTC 2136
Db 1877 TCCAGGATGGGAGGTAGCAATTTGAAATGTAAGAACCATGTGTCCAGACCAAGTCTC 1936
Qy 2137 AGGCCAGCAACTTCTCCCGGACACCGGCTTTCAGCTGTGACCCCAAGCTGGACTGG 2196
Db 1937 CGGCCACGAACTATCTTCAAGAAAGTGGCTCTGACAGTGTGACCTTAAGTCTGG 1996
Qy 2197 ACAGACTGTTCTATGAGATCTGTGCTGCCGACTGTGTGGCCATGGCGTGTGCTAGG 2256
Db 1997 CCCAGACTGCTCAACAGCAATATGTTCTGTGACTGTGGCTTCACACGGCGTTTGCATGG 2056
Qy 2257 GGGCACTGCGCTGCGAGGATGCTGATGGGCGAGCTGCGACAGCGGGCTGCCA 2316
Db 2057 GGGGACGTGTGCTGTGAAGAGGTGGAACGGGCGCCAGCTGTGTAATCAGAGAGCTGCCA 2116
Qy 2317 CCCGCTGTGCCGAGCATGGGACCTGCGCGCAGCGCAAGTGCAGTGCAGCCCTGGCTG 2376
Db 2117 CCCCGCTGTGCCGAGCAGGACCTGCAAGGATGGCAAGTGTGAATGCAAGCCAGGGCTG 2176
Qy 2377 GAATGGCAACACTGCACTGCTCACTATCTGATAGGAGTGT----- 2421
Db 2177 GAATGGAGCACTGCACTATCGCTCACTATTTGGATAAGATAGTTAAAGACAAGATAGG 2236
Qy 2422 ---TAAAGAGGTTTCCCTGGGTTGTGCAATGGCAACGGCAGAGTGTACCTTAGACCTGAA 2478
Db 2237 ATATGAAGGTTTGTCTGCTGTGCAACAGCAATGGAAGATGTACCTGTGACCAAAA 2296
Qy 2479 TGGTGGCACTGCTGCTGCGAGTGTGGCTGAGAGAGCTGGCTGTGACATTCATGGA 2538
Db 2297 TGGCGCAATTGTGTGCCAGCTGGATGGAGAGGAGCTGTGACGTTAGCCATGGA 2356

Qy 2539 GACTGCTGCGGTGACAGAAAGACAATGATGAGATGGCTGTGGTGGATGATGACCC 2598
Db 2357 GACTCTTTGCAAGATAGCAAGGACAATGAAGGGATGAGACTCAATGCTGATGATCC 2416
Qy 2599 TGACTGCTCCCTCCAGCCCTGTGCCATATCAACCCGCTGTGCTTGGTCCCTTAACCC 2658
Db 2417 CGATTGCTGCTACAGAGTTCCTGCCAGATCAGCCCTATTGTGCGGAGCTCCCGATCC 2476
Qy 2659 TCTGACATCATCCAGGAGACAGTCCCTGTGTCAAGCAAGACCTACATCTCTTA 2718
Db 2477 TCAGGACATCATTAGCCAAAGCTTTCAATCGCCTTCTCAGCAAGCTGCCAAATCTTTTA 2536
Qy 2719 TCACCGCATCAAGTTCTCTGCGGAGGAGACAGCAGCACAATAATCCCGGGGAGAACCC 2778
Db 2537 TGAATCGATTCCTTATAGGATCTGATAGCACCCTATGTTATCCTGGAGAAAGTCC 2596
Qy 2779 CTTTGATGAGGCGATGCTTGTGTTATTGCTGGCCAAAGTGTGACATCAGATGGAACCC 2838
Db 2597 TTTCAATAAGAGCTTGCATCTGTCTCATCAGAGGCCAAGTACTGCTGTGATGGAATCC 2656
Qy 2839 CTGTTGTGTGGAACATCAGTTTGTCAATAACCTCTCTTTGGATATACATCAGCAG 2898
Db 2657 ACTTATGAGTAAATGCTCGTTTTTCCATTCACAGAAATGATATACTATTACCG 2716
Qy 2899 GCAAGTGGCAGCTTGTACTTGTGTGACAAATGGCGCATCTCCATCATCTCTCGGTGGA 2958
Db 2717 NCAGAGCGAATGTTTGAATGCTGGGAAATGGTGGGCTCNCATACTTGGTATTGA 2776
Qy 2959 GGGGCACTTTTCATCACAGGAGCACCTGTGGCTGGCCATGGATCGCTCTTGT 3018
Db 2777 AGATCCCCCATTCCTCACTCAGTATCATCTGTGTGATTCATGGAATGTCTTTATGT 2836
Qy 3019 CATGAAACCATCATCATGAGACATGAGGAGATGAGATTCAGCTGTGACTGAGCAA 3078
Db 2837 GATGATACCTTAGTCTAGTGGAGAAAGAGAGATGACATTCACAGCTGTGATCTGAGTG 2896
Qy 3079 TTTTGGCCGCCCAACCCAGTGTCTCTCCATCCCACTGAGCTCTCCGAGCTCTG 3138
Db 2897 ATTGCTGAGCGCAAAATCCCATCATTTGTGTCTATCCCTTTTTCAGATCTTC 2956
Qy 3139 TGCAGAGAAAGGCCCATTTGCCGAAATTCAGGCTTTGCGAGGAGAAATCTATCTC 3198
Db 2957 TCTGAAAGACAGTCCCATCATTTCCCGAAACACAGGTACTCCAGAGGAACTACATTC 3016
Qy 3199 TGGCTCAAGATGAGCTGAGCTAGCTAGCAGCGGACCCCTGGCTACAAATCTGTCT 3258
Db 3017 AGGAACAGATTTGAAACTCTCTACTTGTGATTCAGAGCTGAGGCTATAAGTCACTCT 3076
Qy 3259 GAGGATCAGCTCACCCACCGACCATCCCTTCACTTCACTCATGAGGTGACCTCATGT 3318
Db 3077 CAAGATCACCATGACCCAGCTATTTATTCATTTAATTAAGAGTTTCATCTTATGT 3136
Qy 3319 AGCGTGGAGGGCGGCTCTTCAGGAAGTGGTTTCGCTGAGCGCCCGAGACCTGTCTTATTA 3378
Db 3137 AGCTGTAGTAGGAAGTCTTCCAAAGTGGTTTCTGCTTCAACCAACTTGGCTTATAC 3196
Qy 3379 TTTTATTTGGGCAAGACAGAGCTGTACAAACAGAAAGTGTGTTGGCTTTCAGAACCTT 3438
Db 3197 TTTTATATGGATGAAACAGATGATATTAATCAGAAAGTCTATGTTCTATCTGAACTGT 3256
Qy 3439 TGTTCCTGGTGGTTATGAATATGAATCCTGCCAGATCTAATCTCTGGGAAAAAGAAAC 3498
Db 3257 TGTCTAGTTGATATGATATGATGCTGTGTTGGACCTGACTCTGTGGGAAAGAGGAC 3316
Qy 3499 AACAGTGTGCGAGGCTATGAAATGACCGCTTCAAGCTTGGAGGATGAGGCTGAGCAA 3558
Db 3317 TGCCATTTCTGAGGCTATGAATTTGGATGCGTCCAACATGGGTGGCTGAGCATTAGATAA 3376
Qy 3559 ACATCATGCCCCCAACATTTCAAGTGGTATCTGTGACAAAGGAAATGGGAGAACAGTT 3618
Db 3377 ACATCATGCTGCTGATGTACAGAACGGTATCTGTACAAAGGAAACGGGGAACAGTT 3436

QY 3619 TGTGCTCAGAGCCCTCTGTCAITGGAGCATCATGGGCAATGGGCGCGAGAGCAT 3678
 Db 3437 CATCTCCAGCAGCCCTCCAGTCGTGAGTAGCATCATGGCAATGGCGAGCCGACAT 3496
 QY 3679 CTCCTGCCAGCTGCAACGGCTTGTGACGGCAACAGCTCTGGCCCGAGTGGCCCT 3738
 Db 3497 TTCTGCCGCCAGTTGCAATGCTCAAGCTGATGTAACAAGTTACTGGCCCGCAGTGGCGCT 3556
 QY 3739 CACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTTCACTACATTAAGAGGATCTT 3798
 Db 3557 AGCTTGTGGATCGATGCACTGTGCTAGTAGGCAATTTCACTAGTGGCGGATATT 3616
 QY 3799 CCCCTCTGGAATGTCACCAATCTCTAGACTGAGGAATAAGATTTCAGACATAGTCA 3858
 Db 3617 CCCCTCTGGAATGTAACAAGTGTCTTAGAACTAAGAAATAAGATTTTAGACATAGCAG 3676
 QY 3859 CAGTCCAGCACAAATACTACCTGGCCACAGACCCCATGAGTGGGCGCTCTTCCTTC 3918
 Db 3677 CAACCCAGCTCATAGATACTACCTTTGCAACGGATCCAGTCAACGGAGATCTGACGTTTC 3736
 QY 3919 TGACAGCAACAGCCGGCGCTCTTTAAATCAAGTCCACTGTGTGTGTAAGGACCTTGT 3978
 Db 3737 TGACACAAACCCCGCAGAAATTTATCCGCAAGTCACTTACGGGGCAAAAGACTTGAC 3796
 QY 3979 CAAGNACTCTGAGTGGTGGGGACAGGTGACAGTGCCTCCCTTTGATGACACTCG 4038
 Db 3797 TAAAAATGCAAGTGTGTCGAGGGACAGGGGAGCAATGCCCTTCCGTTTACGAGGCGAG 3856
 QY 4039 CTGGGGGATGGTGGGAAGGCCACAGAAAGCCACACTCACCAATCCAGGGGTATTACAGT 4098
 Db 3857 ATGTGGGATGGAGGAGGCCGTGGAAAGCCCACTCATGATCCCAAGAGATGGCAGT 3916
 QY 4099 GGACAAGTTTGGGCTGATCTACTTCTGTTGATGGACCAATGATCAGAGCATTCGATCAGAA 4158
 Db 3917 TGATAAGAAATGGATTAATCTACTTGTGTGTAAGCAACATGATTAGGAAGTTGACCAAAA 3976
 QY 4159 TGGGATCATCTCCACCTGTCGCTCTAATGATCTCAGATCAGCCGGGACCTCAGCTG 4218
 Db 3977 TGGAAATCATTAATCTTCTGGGCTCTAAGATTTGACTTTCAGCCAGACCTTTAACTTG 4036
 QY 4219 TGATTTCTGATGATATTTCCAGGTAAGACTGGAGTGGCCACAGACTTAGCCATCAA 4278
 Db 4037 TGACACAGCATGCATCAGCCAGGTACGCTCTGGAATGGCCACTGACCTAGCCATTAA 4096
 QY 4279 CCCAATGACACATCACTTTATGTCCTCGACAAATGTCCTGCAAAATCTTGAAA 4338
 Db 4097 CCCTATGATTAATCTCACTTTATGTCCTGGATAATGATGATTTTACAGATCACTGAAA 4156
 QY 4339 CCACAGGTGGCATTTGTGGCGGAGGCCCATGCACTGCGAGTCCCTGGCATTTGACCA 4398
 Db 4157 TCGTCAAGTTGCGATTTGCTGTGACGCCCATGCACTGTGAGTTCCCGAGTGG--A 4213
 QY 4399 CTTCCTGTAAGCAAGTGGCCATCCAGCAACCTGGAGTCAAGCAGCCGCTTTGGGTGT 4458
 Db 4214 ATATCTGTGGGGAAGCAGCGGTGACAGCAACACTGGAATGATGATTTTACAGATCACTGAAA 4273
 QY 4459 TTCAACAATGGGGTCTGTATATTTGCTGAGACTGATGAGAAAAGATCAACCGCATCAG 4518
 Db 4274 GTCTTACAGTGGGTCTGTACATTTAGAACTGATGAGAAAGAAATTAACCGGATAG 4333
 QY 4519 GCAGGTCAACACTAGTGGAGATCTCACTGTTGTGGGGCCCCCAGTGGCTGTGACTG 4578
 Db 4334 GCAGGTCAACAACAGATGGAGAAATCTCTTAGTGGCGGAATACCTTTCAGAGTGTGACTG 4393
 QY 4579 TAAAAATGATGCCAATGATGTTTCTTCTGGAGAGATGGTTATGCCAAGATCAGAA 4638
 Db 4394 CAAAAATGATGCCAATGATGTTTCTTACAGAGTGGAGATGGCTACGCCAAGATGCCAA 4453
 QY 4639 GTTAAATACCCATCTTCTTCTGGCTGTGTGCTGATGGGAGCTTACGTGGCCGACCT 4698
 Db 4454 ACTCAGTGGCCCATCTCTCTGGCTGTCTTCCAGATGGTACACTGTATATTCAGATCT 4513
 QY 4699 TGGGAACATCCGATTCGGTTTATCCGGAGAACAGCCCTTCTCTCAACCCAGAACAT 4758

Db 4514 AGGGAATATCCGATCCGGCTGTGTCAAGAAATGAAGCTTTTACTTAATCTATGAACCT 4573
 QY 4759 GTATGAGCTGTCTTCAACAAATGACAGAGCTCTATCTGTTGATACACCGGCAAGCA 4818
 Db 4574 CTATGAAGTGGCTCTCCAACTGATCAAGAACTCTACATCTTTGACATCAATGCTACTCA 4633
 QY 4819 CTTGTACACCCAAAGCCTGCCACAGGAGCTTACCTGTACAACTTCACTACCTACCTGGGA 4878
 Db 4634 CCAATATCTGTAAGTTAGTCACTGGTGATCTCTTTACATTTAGTACAGCAATGA 4693
 QY 4879 CCGCAGATCACTCATCAGACAAACAATGGCAACATGTAATGTCCCGGAGATC 4938
 Db 4694 CAATGATATTAAGTGTGACAGACAGCAATGGCAACACCTTGAATTTAGACGGGACCC 4753
 QY 4939 TACTGGATGCCCTCTGCTGGTCCAGAGTGGCCAGGTGCTGAGTGGTACCTGGTACCATGGG 4998
 Db 4754 AAATCGATGGCAGTTCGAGTGGTGTCTCTGATTAACCAAGTGATGTTGTAATAGG 4813
 QY 4999 CACCAACAGTGCATCAAGAGTGTGACACACAAAGGACACAGTGGCCATGATGACATA 5058
 Db 4814 AACAAATGATGTTTGAAGGCACTGCTCAAGGACTGGAAATAGTTTGTACTTA 4873
 QY 5059 CCATGGCAATCCGGCTTCTGGCAACCAAAAGCAATGAAGCAAGTGGACAACTTTA 5118
 Db 4874 CCATGGCAATAGTGGCTTTTAGCCACTAAAAGTGAATGAAGTGGACAACTTTT 4933
 QY 5119 TGAGTACCAACAGCTTTGGCCGCTGACAAATGTGACCTTCCCTACTGGCCAGGTGAGCAG 5178
 Db 4934 TGACTATGACAGTGAAGTCTGTGACAAATGTTACGTTTCCACTGGAGTGGTCAAAA 4993
 QY 5179 TTTCCGAAGTGNATCAGACAGTTCAGTGCATGTCCAGTGAAGACCTCCAGC---AAGGA 5235
 Db 4994 CTTGATGGGACATGGCAAGGCTTATCACAGTGGACATTCAGTGCATCTAGCCGAGAAGA 5053
 QY 5236 TGATGTACCAATACCAACCACTGCTGCTCGCTCAGGGCCCTTCTACACACTGCTGCAAGA 5295
 Db 5054 AGATGTAGATCACTTCAAAATCTGTCTCGATGANTCTTTCTACACCATGGTTCAGA 5113
 QY 5296 CCAAGTCCGGAACAGTACTACATCGGGGCCATGGTCTCTTGGCGTGTGCTGGCCAA 5355
 Db 5114 TCAGTTAAAGAAACAGCTACAGATTGTTATGACGGCTCCCTCAGAAATTAATCAGCCAG 5173
 QY 5356 CCGCATGAGGTGGCGTGCAGACTGAGCCCACTTGTGCTGGCTGGCAGCTCAACCCAC 5415
 Db 5174 TGGCTGGAATCACTACACTACCAACAGAGCCGCACTTCTGGCTGGCAGCTTAATCCGAC 5233
 QY 5416 CTTGGGCAAGAGGAATGTCAAGCTGCCATCGAACAGGCCCTCAACCTGGTGGAGTGGCG 5475
 Db 5234 GTTGCCAAAGAAACATGACTTTGCTGGCGAGAACGTTCAAACTTGGTGGATGGAG 5293
 QY 5476 CGAGCGAAAGAGAGGCTCGGGGCGAGTCACTGTCTTTGGGCGCGGCTGGCGGTGCA 5535
 Db 5294 ATTCCGAAAGAGAGCCCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAA 5353
 QY 5536 CAACCGAAATCTCTATCTCTGGAATTTGATCGGTAAACGACAGAGAGATCTATGA 5595
 Db 5354 TGGCAAGAACTCTCTTTCAGTTGATCGAACAAAGACAGAAAGATCTATGA 5413
 QY 5596 TGACACCGCAAGTTCACCTTCGGATTTCTGACGACAGCGGGGCGGCGGCTCTG 5655
 Db 5414 CGACCACTGTAATTTCTACTGAGGATCGCTTACGACAGCTTGGGCAACCGACTCTCTG 5473
 QY 5656 GTCACCCAGCAGAGGCTGATGTTGTCAGTCAAGTCACTACTCCCTGGGGTTCATTCG 5715
 Db 5474 GCTGCAAGCAGAGCTGATGAGCGGCTGATGCGGCTCATGTCACTTATTCACAGTCAATTCG 5533
 QY 5716 TGGCATCCAGAGGGGCTATGTTCTGTAAGAAATGGAATACAGCAGCGGGCGGCTCAC 5775
 Db 5534 CAGCATCCAGCGAGGCCACCTAGCGAGAAAGTATGATGATGACGACAGCGGAGGATCGT 5593
 QY 5776 ATCAGGATCTTCTGATGGAGAGCATGGAGTACATACTTACAGATCCATGGT 5835

5594 GTCTCGGGTCTTTGCTGATGGTAAACATCGAGTTACACATATTTAGAAAAGTCCATGGT 5653 Db
5936 GCTGCTACTACAGCCAGAGGCGATATCTTTGAGTTTCGACAGAAATGACCGCTCTC 5895 QY
5654 TCCTTCTGCTTATAGCCAGCGGAGTACATCTTCGAATACGATATGGGACCGCTGTC 5713 Db
5996 TTCGTGACGATGCCCAAGCTGGCGCGGAGACACTAGACACCATCGCTCACTGGGGCTA 5955 QY
5714 TGCCATCACCATGCCAGTGTGGCTCGCCACACCATCGATCCATTTGGCTA 5773 Db
5956 CTACAGAAACATCTATCAGCGCCCTGAGGGCAATGCTCTAGTCAGTANACAGACTTCTACTGA 6015 QY
5774 CTACCGCAACATATACAAACCCCGGAAAGCAAGCCCTCCATCATCAGGACTACACGA 5833 Db
6016 GGATGGGCACCTCTTACACACTTCTACCTGGGCACCTGGCGCAGGGTGATATACAAAGTA 6075 QY
5834 GGAAGGGCTGCTTCTACAAACAGCTTCTTTGGGTACAGTTCGGAGGGTCTTATTTCAAATA 5893 Db
6076 TGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACACCAAGGTCTAGTTTCACCTA 6135 QY
5894 CAGAAGGCAGACTAGGCTCTCAGAAAATTTTATATGATAGACAAAGATCAGTTTACCTA 5953 Db
6136 TGACGAGACGCGAGCGATGCTGAAGACCATCAACCTACAGAAAGAGGCTTCACTTGCAC 6195 QY
5954 TGATGAACACAGCAGGAGTCTTAAAGACAGTAAACCTCCAGAGTATGGTTTTATTTGCAC 6013 Db
6196 CATCGGCTACCGTACAGATTGGGCGCCCTGATGACCGACAGATCTTCGCGTTCACCTAGGA 6255 QY
6014 CATTAGATACAGCAATTTGCTCCCTGATTTGACAGGCAGATTTTCGCTTTAGTGAAGA 6073 Db
6256 AGGCATGTGTAACGCCGCTTTTGAATCAACTATGACAAACAGCTTCGGGTGACAGCAT 6315 QY
6074 TGGGATGTTAAATCAAGATTGACTATAGCTATGACAAACAGCTTTCGAGTGACCAAGCAT 6133 Db
6316 GCAGGCTGTATCAACGAGACCCCACTGCGCATTTGATGCTCTATCGCTATGATGATGTGC 6375 QY
6134 GCAGGCTGTATCAATGAACCGCAGCTGCTTATGATCTGATCAGTTTGAATGACATTTTC 6193 Db
6376 AGGCAAGACAGCAGGATTTGGGAAGTTGGTGCTCATTTACTPATGACATTAACCAAGTATCAT 6435 QY
6194 TGGCAAAAGTTGAGCAGTTTGGAAAGTTTGGAGTTTATATATGATATTAACACAGATCAT 6253 Db
6436 CACCACAGCTGTATGACCCACACCAAGCATTTTGTATGCTATATGGCAGGATGAGGAAGT 6495 QY
6254 TTCTACAGCTGTATGACCTATAGCAAGCATTTTGTATGCTCATGGCCGTATCAAGGAGAT 6313 Db
6496 GCAGTATGAGATCTTCCGCTCGCTCATGTACTGTGATGACCTCCAGTATGATTAACATGGG 6555 QY
6314 TCATATGAGATATTCAGTTCGCTCATGTAAGTATGATTAACATGATGATTAACATGGG 6373 Db
6556 GCGAGTATGAGAAAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCTACTCCTA 6615 QY
6374 TCGGTAACCAAGAGAGAGATTAATAATAGGGCCCTTTTGCCAAACACCAAAATATGCTTA 6433 Db
6616 TGAGTATGATGTGACGGCCAGCTGCAGACAGTCTCCATCAATGACAAAGCCACTCTGGCG 6675 QY
6434 TGARTATGATGTTGATGACAGCTCCAAACAGTTTACCTCAATGABRAGATAATGTGGCG 6493 Db
6676 CTACAGCTACGACCTCAATGGGAACCTGCACTTACTAGCCCTGGGAACAGTGCACGGCT 6735 QY
6494 GTACAACTACGATCTGAATGAAACCTCCATTTACTGAACCCCAAGTAACAGTGGCGGCTCT 6553 Db
6736 CACACCACTACGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGACGTGCAATACAA 6795 QY
6554 GACACCCCTTCGCTATGACCTCGGAGACAGATCACTCGACTGGGTGATGTTCAATATCG 6613 Db
6796 GATGATGAGGATGGCTTCCTGAGGCAGCGGGCGGTGATATCTTTTGATGTAACAATCAGC 6855 QY
6614 GTTGGATGAGATGTTTCTTACGTCAAAGGGGACGGAATCTTTGATATAGCTCCAA 6673 Db
6856 TGGCTGTCTATCAAGGGCTTACAACCGGGCTGGCAGCTGGAGTGTGAGTACCGGTACGA 6915 QY
6674 GGGGCTTCTAAGTCAAGTAAAGGGCAGTGGCTGGAAGTCAAGTATCTACCGTTATGA 6733 Db

6916 TGGCCTGGGGCGCGGTGCTCCAGCAAGAGAGCCACACCCACCTGCAAGTTCTTCTA 6975 QY
6734 CGGCTTGGGAGGCGGTGTTTCTAGCAAAACAGTCTAGGACAGCACTGCGATTTTTTA 6793 Db
6976 TCGACACTGACAAACCCACCAAGGTCAACCACTGTACAACTCTCAGCTCTGAGAT 7035 QY
6794 TGCTGACTTAACCTTATCCACTAGGATTAATCATCTCTACCAACCATTCGAGTTCAGAAAT 6853 Db
7036 CACCTCCCTTACTAGGACTTGCAGAGGACACTCTTTGCTTGGAGCTGAGCAGTGGTGA 7095 QY
6854 TACCTCCCTGTATTTATGATCTCCAAGGACATCTTTTGGCCATGGAAATCAGCAGTGGGA 6913 Db
7096 TGAGTTTTACATAGCTTGTGACACATCGGACCCCTCTTGTCTGTTTTAGTGGAAACAGG 7155 QY
6914 TGAATTTCTATAATTGCATCGGATAACACAGGAGACCACTGGCTGTGTTTCAAGTAGCAATGG 6973 Db
7156 TTTGATGATCAAGCAAACTCTGTACAGCCCTATGGGAGATCTACATGATACCAACCC 7215 QY
6974 GCTTATGCTGAACAGATTCAGTACACTGCATATGGGAAATCTATTTTGTACTTAATAT 7033 Db
7216 CAACTTTTCAGATCATCATAGGTACACCTGGTGGCTCTATGATCCACTCACCAAGCTTGT 7275 QY
7034 TGACTTTCACTGGTAAATTTGGATTTTCAATGGTGGCTGTATGACCCACTCACCAATTAAT 7093 Db
7276 CCACATGGGCGCGGAGATTTATGATGCTGGCCGAGCGCTGGACTAGCCAGACACGA 7335 QY
7094 CCACITTTGGAGAAAGAGATTAATGACATTTTGGCAGACGCTGGACAAACCTGCATAGA 7153 Db
7336 GCTGTGGAGACACCTTAGTAGCAGCAAGCTCATGCTCTTAACTCTATATGTTCAAAAA 7395 QY
7154 AATCTGGA---AAGAAATTTGGGAAGGACCCAGCTCTTTTAACTTTGTACTATGTTAGGAA 7210 Db
7396 CAAACACCCCATCAGCAACTCCAGGACATCAAGTCTTCATGACAGATGTTTAAACAGCTG 7455 QY
7211 TAAACACCCCTGCAAGCAAAATCCATGACGTGAAGATTAACATCAGAGATGTTTAAACAGCTG 7270 Db
7456 GCTGCTCACTTTTGGATTTCCAGCTACACAACTGATCCCTGGTATCCCAACACAGACAT 7515 QY
7271 GCTGTGATCATTTGGTTTCCATCTGCACAACTGCTATCTCGATTCCTGTTCCCTCCCAATTT 7330 Db
7516 GGATGTCATGGAACCCCTCCTACGAGCTCATCCACACACAGATGAAACACGAGAGTGGGA 7575 QY
7331 TGATTTAAACAGAACTTCTTACGA-----ACITTTGAAGAGTTCAGCAGTGGGA 7378 Db
7576 CAAACAGAGTCTATCTCGGGGTACAGTGTGAAGTACAGAAAGCAGCTCAAGGCTTTTGT 7635 QY
7379 TGATATACCGCCCATCTTCGGAGTCCAGCAGCAAGTGGCGCGGCGAGGCCCAAGGCTTCTCT 7438 Db
7636 CACCTTAGAACGCTTTGACAGCTCTATGGCTCCACAATCACCAGCTGCCAGGCTCC 7695 QY
7439 GTGCTGGGAAGATGCGCGAGGTGCAGGTGAGCGCGCGCGCGCGCGC---GGCGCGCA 7495 Db
7696 AAAGACCAAGAGTTTGCATCCAGCGGCTCAGTCTTTTGGCAAGGGGTCAAGTTTGGCTTT 7755 QY
7496 GTCTGGCTGTGTTTCGCCACCGTCAAGTCTGATCGGCAAGGGCGTCTATGCTGCCCT 7555 Db
7756 GAAGGATGGCGGAGTACACACAGACATCATCAGTGTGGCCAATGAGGATGGCGAAGGT 7815 QY
7556 CAGCCAGGGCGCGTGCAGACCAACCTGTCTAACTCGCCACAGAGATGATCAAGGT 7615 Db
7816 TGCTGCTCATCTTGAACATGCTCCACTACCTAGAGAACTTGCACCTTCAACATTTGATGGGT 7875 QY
7616 GCGCGCGTCTCAACAAACGCTTCTACCTGGAGAACCTTGCATTCACCATCGAGGGCAA 7675 Db
7876 GATACCCATTTACTTTGTGAACACAGGACCTTCAGAAAGTGTGACCTGGCCATCTGGGCT 7935 QY
7676 GGCACGCGCTACTTCTCAAGACCAACCGCCGAGAGCGCTGGGACCGCTGGGCTGGGT 7735 Db
7936 CAGTGGGGGCGCGCAACCTGGAGAAATGGGGTCAACGTCACCTGTCTCCAGATCAACAC 7995 QY
7736 GACCAGCGCGCGCAAGGCGCTGGAGAAACGGCATCAACGTGACGGTGTCCAGTCCACCAC 7795 Db

QY 7996 AGTACTTAATGTCAGGAGTACGCTACACAGACATCCAGCTCCAGTACGGGGCACTGTG 8055
 Db 7796 GGTGTGTAACGCGAGGACCGCAGGTTCCGGACGTGGAGATGCGGTTCGGCGCTGGC 7855
 QY 8056 CTTGAACACAGCTACGGGACAACTTTGGATGAGGAGGACGCGGTCTCTGGAGCTGGC 8115
 Db 7856 GCTGCACGTGCGTACGCGATGACCTCGACGAGGAGGAGGCGCATCTCTGGAGCAGGC 7915
 QY 8116 CCGGAGAGAGCGTGCAGGAGCGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8175
 Db 7916 CCGGAGAGAGCGTGCAGGAGCGTGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7975
 QY 8176 GGAAGGCTCGGGGCTCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8235
 Db 7976 GGAAGGCTCGGGGCTCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8035
 QY 8236 GGAAGGCTCGGGGCTCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8295
 Db 8036 GGAAGGCTCGGGGCTCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8095
 QY 8296 GGAAGGCTCGGGGCTCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8345
 Db 8096 GGAAGGCTCGGGGCTCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 8445
 RESULT 7
 ID ABQ82346
 AC ABQ82346
 DT 17-DEC-2002 (first entry)
 DE Human NOV15d encoding cDNA SEQ ID NO:41.
 KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
 KW cerebroprotective; neurotropic; antidiabetic; antiinflammatory; fungicide;
 KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
 KW antianemic; antibacterial; protozoacide; antihelminthic; gene therapy;
 KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
 KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
 KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
 KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
 KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
 KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
 KW rheumatoid arthritis; gene; chromosome 4; ss.
 OS Homo sapiens.
 XX
 XX
 XX
 PH Key Location/Qualifiers
 FT 299..8140
 FT /*tag= a
 FT /product= "NOV15d"
 FN WO200262999-A2.
 XX
 XX
 XX
 PD 15-AUG-2002.
 XX
 PF 31-DEC-2001; 2001WO-US049976.
 XX
 PR 29-DEC-2000; 2000US-0258928P.
 PR 02-JAN-2001; 2001US-0259415P.
 PR 04-JAN-2001; 2001US-0259785P.
 PR 20-FEB-2001; 2001US-0263814P.
 PR 09-MAR-2001; 2001US-0279863P.
 PR 29-MAR-2001; 2001US-0279832P.
 PR 13-APR-2001; 2001US-0279833P.
 PR 18-APR-2001; 2001US-0283889P.
 PR 25-APR-2001; 2001US-0284477P.
 PR 29-MAY-2001; 2001US-0294080P.
 PR 16-AUG-2001; 2001US-0312915P.
 PR 17-AUG-2001; 2001US-0313325P.

PR 17-SEP-2001; 2001US-0322699P.
 XX 26-NOV-2001; 2001US-0333350P.
 PA (CURA-) CURAGEN CORP.
 XX
 XX
 PI Spytek KA, Li L, Wolenc AR, Vernet CM, Eisen A, Liu X;
 PI Malyankar U, Shinkens RA, Tchervnev VT, Spaderna SK, Gorman L;
 PI Kekuda R, Paturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
 PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
 PI Gunther E, Smithson G, Millet I, Macdougall JR;
 XX WPI; 2002-732706/79.
 DR P-PSDB; ABP53589.
 XX
 XX
 PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
 PT associated disorders, such as cancers, neurological disorders, disorders
 PT of vesicular transport, gastrointestinal disorders, and autoimmune
 PT diseases.
 XX
 PS Claim 8; Page 123-125; 444pp; English.
 XX
 CC The present invention describes novel human proteins designated NOVX,
 CC where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
 CC cytostatic, anticonvulsant, cerebroprotective, neurotropic, cardiovascular,
 CC antidiabetic, antiinflammatory, antihelminthic, antianemic, virucide,
 CC immunosuppressive, antiallergic, antianemic, antibacterial, fungicide,
 CC protozoacide and antihelminthic activities, and can be used in gene
 CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
 CC manufacture of a medicament for treating a syndrome associated with a
 CC human disease selected from NOVX-associated disorder, such as cancers
 CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
 CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
 CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
 CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
 CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
 CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
 CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
 CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
 CC and protozoal infections. The NOVX proteins can be used as immunogens to
 CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
 CC used in chromosome mapping, identifying individuals from minute
 CC biological samples (tissue typing), and in forensic identification of a
 CC biological sample. The present sequence encodes human NOV15d, which is
 CC located on chromosome 4
 XX
 SQ Sequence 8487 BP; 2306 A; 2029 C; 2148 G; 2001 T; 0 U; 3 Other;
 Query Match 40.7%; Score 3399.8; DB 6; Length 8487;
 Best Local Similarity 67.0%; Pred. No. 0;
 Matches 5007; Conservative 2; Mismatches 2361; Indels 102; Gaps 9;
 QY 880 CGGCACCTTCCTTCAGGCTGG---AGGCACCTCCCGCTCTCTGCACCATCACC 936
 Db 766 CAGGCATTTCTATTCAAAACAGGACAGGATACAGCCACTGTTAGTACTGCAACCCC 825
 QY 937 AGGTACCCACTGAGCTCCAGCAGTGTACTCTCTCCGCCCGCCAGCCCTGCCCGGAG 996
 Db 826 AGGATACAAATGGCATCTGGCTCTGTTTATTCACCACTACTCGGCCACTACCTAGAAA 885
 QY 997 CACCTTCGCCCGCGCGCTTTAACCTCAAGAACCTCCAGTACTGTAACCTGGAAGTG 1056
 Db 886 CACCTTATCAAGAGTGGCTTTAAATTCAGAACTCTTCAAGTACTGTAGTGGAAATG 945
 QY 1057 CGCAGCCTGAGCGCCATCGTCACTCTCAGGCACCTCTGCTATCTGTGTGCTACTTTGT 1116
 Db 946 CACTGCACTGTGTGCGCTAGGGGTCTCGGTGCTCTCGTCAATCTCTGCTTTATTTAT 1005
 QY 1117 GGCATGACCTGTTTGGCTTAACTGACCTGACCTGACCTGAGGGGAGAGTATGA 1176
 Db 1006 AGCAATGCACTCTTTTGGCTTCAACTGGCAGCTTACAGCAGCTGAAATGACACATTGA 1065
 QY 1177 GATCAGGAGGACACAGCCAGTGTGGCTGTGTGCCAACCGAGCTCTCCCTATACCCCTC 1236

Db 1066 GA-----ATGAAAAGTGAATCTTGATACCATGCCAACAAACACACTGTGTCATTACCTTC 1119
QY 1237 AGGGGCACTGGCTTAGAGACCCCTGACAGGAAGGCAAGAACACACAGGAAGGCC 1296
Db 1120 TGGAGACA-----ATGAAATAATT 1137
QY 1297 CAGTAGTTTCTTTCCAGAGGACAGTTCATAGATTCTGGAGAAATGATGTGGGAAGCG 1356
Db 1138 AGGTGGATTACGCAAGAAAATAACACATAGATTCCGGAGAACTTGATATTGGCCGAAG 1197
QY 1357 AGCCTCCAGAAAGATTCCTCTGSCACTTCTGAGATCTCAAGTGTTCATAGACCATCC 1416
Db 1198 AGCAATTCAAGAGATTCTCCCGGATCTCTGAGATCAAGCTCTCAATGATACGCC 1257
QY 1417 TGTGATCTGAAATTCATGTGTCTCTGGGAAAGCGACGCCCTGTTGGCAATTTATGGCAG 1476
Db 1258 ACAGTTCTTAATCAATATCTCTCTCAGAAAGATGCAATTGATGGAGTATATGSCCG 1317
QY 1477 AAAAGGCTCCCTCTCTCAATACACAGTTCATCTTTGTGGAGCTGTGATGGCAGAG 1536
Db 1318 GAAAGGCTTACCGCTTCCCATACTCAGTATGACTTCGTGGAGCTCCTGGATGGCAGAG 1377
QY 1537 GCTCTTAACCCAGAGGCGCGAGCTAGAGGGACCCGCGCCAGTCTCGGGAACTGT 1596
Db 1378 GCTGATGCCAGAGCAGCGAGCCCTCTTGAGACGAGAGCCCGGCGGAGGCGAG 1437
QY 1597 GCCCCCTCCAGCCATGAGACAGGCTTCATCCAGTATTTGGATTCAGGAATCTGGCACTT 1656
Db 1438 ATCCGTACGCTTCATGAGCGCGCTTTATCCAGTACTTGGATTCTGGAATCTGGCATCT 1497
QY 1657 GGCCTTTTACATGACGAAAGGATCAGAGTGGTTTCCCTTCTCACCCTGCCATTGA 1716
Db 1498 GGCCTTTTATATGATGGGAAAAATGACAGAGAGGTGTCTTTTATACCATTTGTTAAGA 1557
QY 1717 GTCGCTGATAACTCCGCCAGCACTGCTATGGCAATGGTGACTGCATCTCTGGGACCTG 1776
Db 1558 GTCTGTGTGGAAATGTCCTCCGAAATTCATGGAAATGGAAATGGGTTCTTCGAACTTG 1617
QY 1777 CCACTGCTTCTGGGTTCTCTGGGCCCGACCTGTGGCAGAGCTCTCTGCCCGCTGCTG 1836
Db 1618 CCAATGTTTTCAGGATTTCTGGGTCGGGATTTGTTCAAGAGCCGCTGTCCAGTGTATG 1677
QY 1837 TAGCGGAAATGCCAAATACATGAAGGAGATGCTTTGTCACAGTGGCTGGAAAGGCGC 1896
Db 1678 TAGTGGCAAGGGCAGTACTCCAGGGCGCTGCTGTCTTTCAGCGCTGGAGGGCAC 1737
QY 1897 TGAGTGCAGTGGCCCAACCAAGTGTATCGATGTGGCTGTGACAGCAACCAATGGCACTG 1956
Db 1738 CGAGTGTGATGTGCCGACTACCCAGTGTATTGACCCACAGTGTGGGGTCTGTGGGATTTG 1797
QY 1957 CATCAGGGCACTGCATCTGCAACCTGGCTACAGGGCGAGAGCTGTGAGGAAGTGA 2016
Db 1798 TATCATGGGCTCTCTGTGCTTGCATCTCAGGATACAAAGAGAAAGTTGTGAAGAGCTGA 1857
QY 2017 CTGCATGACCCCAATGTTTCAGGCGGGGTGTCTGCGTGAGAGCGGAATCCATGCTT 2076
Db 1858 CTGTATAGACCTCGGTGTCTTAATCATGTGTGTATCCAGGGGAATGTCACTGCAG 1917
QY 2077 TGTGGATGGGAGGCAACCACTGAGACCCCGAGGCCCACTGCTTAGACCAAGTTC 2136
Db 1918 TCCAGGATGGGAGTAGCAATTTGAAATACTGAAGACCAATGTGTCCAGACCAAGTGTCT 1977
QY 2137 AGGCCACGGAACCTTCTCCCGACACCGGGCTTTGACGTGTGACCAAGCTGAGCTGG 2196
Db 1978 CGGCCACGGAACGTATCTTCAGAAAGTGGCTCTCTGCAGTGTGACCTTAAGTGGCTGG 2037
QY 2197 ACACGACTGTTCTATCGAGATCTGTGCTGCCACTGTGTGGGCCCAATGGCGTGTGGTAGG 2256
Db 2038 CCCAGACTGCTCAAAACGAAATATGTTCTGTGACTGTGGCTCACACGGCGTTTGCATGGG 2097
QY 2257 GGGCACTGCCCTGGAGGATGGCTGGATGGGGGAGCTGTCGACACGCGGCTGCCA 2316
Db 2098 GGGGACGTGTCTGTGAAGAAAGGCTTGGACGGGCGCAACCTGTATACAGAGCGCTGCCA 2157

QY 2317 CCCGCGCTGTGCGAGCTGGACCTGCGCGACGCAAGTCCGAGTGCAGCCTGGGTG 2376
Db 2158 CCCCGCTGTGCGGAGCAGGACCTGCAAGGATGGCAAGTGTGAATGCAGCATGGGTG 2217
QY 2377 GAATGGGAAACTGCAACCATCGCTCACTATCTGATAGGTAGTAAAGAGGTTGCC 2436
Db 2218 GAATGGAGAGCACTCAGCTATC-----GAGGGTGTCTC 2250
QY 2437 TGGGTTGTCCAATGGCAACGCGAGATGTACCTTTAGACCTGAAATGGTTGGCACTGGCTG 2496
Db 2251 TGGTCTGTGCAACAGCAGCAATGGAAGATGTACCTTGACCAAAATGGCTGSCATTTGTGTG 2310
QY 2497 CAGCTGGCTGGAGAGGAGTGGCTGTGACACTTCATCTGAGAGACTGCTGGGTGACAG 2556
Db 2311 CCAGCCTGATGGAGAGGAGAGGCTGTGACGTAGCCATGGAGACTCTTTTGCACAGATAG 2370
QY 2557 CAAAGCAATGATGGAGATGGCTGTGTGACTGTGATGACCTGACTGTGCTCAGCC 2616
Db 2371 CAAAGCAATGAGAGAGATGGACTCATTTGACTGTGATGATCCGATTCGCTGCTACAGAG 2430
QY 2617 CCTGTGCCATATCAACCCGCTGTGCTTGGCTCCCTTAACCTCTGGACATCATCCAGGA 2676
Db 2431 TTTCTCCAGAAATCAGCCCTATTGTGGGAGCTGCGGATCCTCAGGACATCATTAGCCA 2490
QY 2677 GACACAGTCCCTGTGTCTACAGAGAACCTAGACTCTTCTATGACCGCATCAAGTCTCT 2736
Db 2491 AAGCTTTCAATCGCTTCTCAGCAGCTGCCAAATCCCTTTTATGATCGAATCAGTTTCT 2550
QY 2737 CGTGGCAGGGCAGCAGCAGCATATCCCGGGAGAACCCCTTTTGTATGGAGGCACTG 2796
Db 2551 TATAGATCTGATAGCACCCATGTTATACCTGGAGAAAGTCTTTTCAATAAGAGCCTTG 2610
QY 2797 TTTGTATTTCGTGGCCAAAGTATGACATCAGATGGAACCCCTCGTTGGTGTGAACAT 2856
Db 2611 ATCTGTCACTCAGAGGCCAAAGTACTGACTGCTGATGGAACCTCCTATTATTTGGAGTAAATGT 2670
QY 2857 CAGTTTGTCAATAAACCCCTCTTTTGGATATACAATCAGCAGGCAAGATGGCAGCTTTGA 2916
Db 2671 CTCGTTTTTCCATTACCCAGATATGATATACTATTACCCGCGCAGGACGAATGTTGA 2730
QY 2917 CTTGTGACAAATGGCGGATCTCCATCATTCCTGCGGTTTCGAGCGGGGCACTTTTCATCAC 2976
Db 2731 CTTGGTGGAAATGTGTGGGCGCTCTCTAACTTTGGTATTGAAACGATCCCATTCCTCAC 2790
QY 2977 ACAGGACACACCTGTGCTGCCATGGGATCGCTTTTGTCTATGCAAGAACCATCATCAT 3036
Db 2791 TCAGTATCATCTGTGTGGATCCATGGAATGTCTTTTATGTGATGATACCTTAGTCTCAT 2850
QY 3037 GAGACATGAGGGAATGAGATTTCCAGCTGTGACCTGAGCAATTTTCCCGGCCCAACCC 3096
Db 2851 GGAGAAAGAGAAATGACATTCAGCTGTGATCTGAGTGGATTCTGTAGGCAAAATCC 2910
QY 3097 AGTCGTCTCTCAATCCCACTGACGTCTCTCGCCAGCTCTGTGCGAGAGAAAGGCCCAT 3156
Db 2911 CATCATTTGTCTATCCTTTATCCACTTTTATCCACTTTTTCAGATCTTCTCTGAGAGAGTCCCAT 2970
QY 3157 TGTGCGGAAATTCAGCTTTGAGGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCT 3216
Db 2971 CATTCGGAACACAGGTACTCCACGAGGAACTCAATTCAGAGAACAGATTGAACT 3030
QY 3217 GAGCTACCTGAGCAGCGGACCCCTGGCTACAAATCTGTCTGAGGATCAGCCTCACCCA 3276
Db 3031 CTCTACTTGTGATTCAGAGCTCAGGGTATAAGTCAGTTCTCAAGATCACCATGACCCA 3090
QY 3277 CCGACCATCCCTTCAACCTCATGAGGTGACCTCATGGTAGCGGTGAGGGCGCGCT 3336
Db 3091 GTCTATTATTCCATTAAATTAATGAGGTTTCTTTATGTGATGCTGTAGTAGAAGACT 3150
QY 3337 CTTTCAGGAAGTGTTCGCTGCAGCCCCAGACCTGTCTCTATTATTTTCTATTGGGACAGAC 3396
Db 3151 CTTCCAAAGTGGTTTCTTCCCTTCCAAACTTGGCTATACCTTTTATATATGGATAAAC 3210

QY 3397 AGACGTCTACACCAAGAGGTGTTGGGCTTTCAGAACGCTTTGTTCCGTGGGTATGA 3456
DB |||||
QY 3211 AGATGATATAATCAGAAGTCTATGGTCTATCTGAAGCTGTGTGTCAGTTGGATATGA 3270
DB |||||
QY 3457 ATATGAATCCTGCCAGATCTAATCCTGTGGGAAAAAAGAACACAGAGTGTGCAGGGCTA 3516
DB |||||
QY 3271 GTATGAGTCGTGTTGGACCTGACTCTGTGGGAAAAGAGGACTGCCATCTCTGAGGGCTA 3330
DB |||||
QY 3517 TGAATATGACGGCTCAAGCTGTGGAGGATGGAGCCCTAGACAAAATCATGTCCTCAACAT 3576
DB |||||
QY 3331 TGAATTTGGATGCGTCCAAATCGGTGGCTGGACATTTAGATAAATCATCACGTCTGGATGT 3390
DB |||||
QY 3577 TCAAAAGTGTATCTCTGCAAAAGGAATGGGAGAACAGTCTGTGTCTCAGAGCCTCC 3636
DB |||||
QY 3391 ACAGAACGGTATCTGTACAGGGNAACGGGAAACCGATTCTCTCCAGAGCCTCC 3450
DB |||||
QY 3637 TGTCAATGGGAGCATCATGGCAATGGGCGCGGAGAACATCTCTGCCCCAGCTGCAA 3696
DB |||||
QY 3451 AGTCGTGATGATCATGGGCAATGGGCGAAGGCGAGCATTTCTCTGCCCACTGCCAA 3510
DB |||||
QY 3697 CGGCCCTTCTGACGCGCAACAGCTCTGGCCCGAGTGGCCCTCACCTGTGGCTCTGACGG 3756
DB |||||
QY 3511 TGGTCAAGCTGATGTATAAAGTTACTGGCCCGAGTGGCGCTAGCTTTGTGGGATCGATGG 3570
DB |||||
QY 3757 GAGCCTCTATGTGGTGTATTTCAACTACATTAGAAGGATCTTCCCTCTGGAAATGTCA 3816
DB |||||
QY 3571 CAGTCTGTACGTAGGCGATTTCAACTACGTGGCGGATATTCCTCTCGAATGTAA 3630
DB |||||
QY 3817 CAAATCTTAGAGTGGAGTAATAAGATTTAGACATAGTGCACAGTCCAGGACACAAATA 3876
DB |||||
QY 3631 AAGTCTTTAGAACTAAGAAATAAGATTTTAGACATAGCAGCAACCCAGCTCATAGATA 3690
DB |||||
QY 3877 CTACCTGCGCCACAGACCCCATGAGTGGGCGGTCTCTCTTCTGACAGCAACAGCGGCG 3936
DB |||||
QY 3691 CTACCTTGCACAGGATCCAGTCAAGGAGATCTGTACGTTCTGACACAAACCCGAG 3750
DB |||||
QY 3937 GGCTTTTAAATCAAGTCCACTGTGGTGTGAAGGACCTTGTCAAGAACTCTGAGTGGT 3996
DB |||||
QY 3751 AATTTATCGCCCAAGTCACTTACGGGGGCAAAAGACTTGACTAAAAATGCAGAACTGT 3810
DB |||||
QY 3997 TGGGGGACAGTGCACAGTCCCTCCCTTTGATGACACTCGCTCGGGGATGGTGGAA 4056
DB |||||
QY 3811 CGAGGACAGGAGCAATGCCCTTCCGTTTGAAGGCGAGATGTGGGGATGGAGGAA 3870
DB |||||
QY 4057 GGCACAGAGCCACACTCACCAATCCAGGGGTATTACAGTGGACAAAGTTGGGCTGTAT 4116
DB |||||
QY 3871 GGCGGTGAAGCCACACTCATGATGATCCAAAGGAATGGCAGTTGATGAAGAATGATTAAT 3930
DB |||||
QY 4117 CTACTTCGTGATGGCACCATTGATCAGACGATCGATCAGATGGGATCATCTCCACCT 4176
DB |||||
QY 3931 CTACTTTGATGGAACCATGATTTAGAAAGTTGACCAAAATGGAATCATATCAACTCT 3990
DB |||||
QY 4177 GCTGGCTCTAATGATCTCAATCAGCGCGGCCACTCAGCTGTGATTTCTGTCAATGATAT 4236
DB |||||
QY 3991 TCTGGCTCTAACGATTTGACTTCAGCGAGACCTTTAACTTTGACACAGCATGCACAT 4050
DB |||||
QY 4237 TTCCAGGTAGACTGGAGTGGCCCAAGACTTAGCCATCAACCCCAATGACAACTCACT 4296
DB |||||
QY 4051 CAGCAGGTGATCTGGAATGGCCCACTGACCTAGCCATTAACCCCTATGGATTAACCTCAT 4110
DB |||||
QY 4297 TTATGTCTCGAACCAATGTGTCCTGCAAACTCTCTGAAACCAACAGGTGGCATTGT 4356
DB |||||
QY 4111 TTATGTCTGATTAATATGTAGTTTACAGATCACTGAAAATGCTCAAGTTCCGATTCG 4170
DB |||||
QY 4357 CGCGGAGGCCCATGCACTGCCAGGTCCCTGGCAATGACCACTCTCTGTAAAGCAAGGT 4416
DB |||||
QY 4171 TGTGTGACGGCCCATGCACTGTCAAGTTCGCGAGTGG---AATATCCTGTGGGGAAGCA 4227
DB |||||
QY 4417 GGCCATCCAGCAACCTGGAGTCAGCCAGCGCTTTGGCTGTTTACAAATGGGCTCT 4476
DB |||||
QY 4228 CGCGGTGAGCAACACTGGAAATCAGCCACTGCGCAATTTGCTGTCTACAGTGGGCTCT 4287
DB |||||
QY 4477 GTATATTTCTGAGACTGATGAGAAAAAGATCAACCGCATCAGGCGATCCACTAGTGG 4536
DB |||||

DB 4288 GTACATTACTGAAACTGATGAGAAGAAATTAACCGGATTAAGCGAGTGCACAAAGATGG 4347
QY |||||
QY 4537 AGAGATCTCACCTCGTGTGGGGCCCCCAGTGGCTGTGACTGTAAATATGATGCCAACTG 4596
DB |||||
DB 4348 AGAAATCTCTTAGTGGCGGGAATACCTTCAGAGTGTGACTGCAAAATATGATGCCAACTG 4407
QY |||||
QY 4597 TGATTTGTTTTCTGGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATAACCCATCTTC 4656
DB |||||
DB 4408 TGACTCTTACCAGAGTGGAGATGGCTAGCCCAAGGATGCCAAACTCAGTGCCCCATCCTC 4467
QY |||||
QY 4657 CTTGGCTGTGTGCTGATGGGAGCTCTACGTCGCGGACCTTTGGGAAACATCCGAATTCG 4716
DB |||||
DB 4468 CTTGGCTGTTCTCAGATGGTACACTGTATATTGAGATCTAGGGAATATCCGATTCG 4527
QY |||||
QY 4717 GTTTATCCGGAAAGAAACAGCCCTTTCTCAACACCCAGAACATGTATGAGCTGTCTTACC 4776
DB |||||
DB 4528 GGCTGTGTCAAAAGAAATAAGCCCTTTACTTAACTCTATGAACCTCTATGAAGTTGCGTCTC 4587
QY |||||
QY 4777 AATTGACAGGAGCTCTATCTGTTTGATACACCGGCAAGCCTGTACCCAAAGCCT 4836
DB |||||
DB 4588 AACTGATCAAGAACTCTACATCTTTGACATCAATGGTTACTCACCATATATCTGTAAGTT 4647
QY |||||
QY 4837 GCCCAGGAGACTACTGTACAACTTCACCTTACACTGGGACGGCGACATCACACTCAT 4896
DB |||||
DB 4648 AGTCACTGGTGATTACCTTTTACATTTAGCTACAGCAATGACAAATGATTTACTGCTGT 4707
QY |||||
QY 4897 CACAGAAACAAATGGCAACATGGTAAATGTCCGCGAGACTCTACTGGGATGCCCTCTG 4956
DB |||||
DB 4708 GACAGACGCAATGGCAACCCCTTAGAATTTAGACGGGACCCAAATCGCATGCCAGTTTCG 4767
QY |||||
QY 4957 GCTGTGTGTCAGATGGCCAGTGTACTGGGTGACCATGGGACCAACAGTGCACCTCAA 5016
DB |||||
DB 4768 AGTGTGTCTCTGATTAACCAAGTGTATGTTGACATAGGAACAAATGATGTTGAA 4827
QY |||||
QY 5017 GAGTGTGACCAACAGGACAGAGTTGGCCATGATGACATACCATGGCAATTCGGGCT 5076
DB |||||
DB 4828 AGGCATGACTGCTCAAGGACTGGAATAGTTTGTGTTTACTTACCATGGCAATGAGGCT 4887
QY |||||
QY 5077 TCTGCGCAACCAAAAGCAATGAACCGATGGACACATTTTATGATGACGACGCTTTGG 5136
DB |||||
DB 4888 TTTAGCCACTAAAGTGTGAACCTGGATGGACAAACGTTTTTGTACTATGACGTGAAGG 4947
QY |||||
QY 5137 CGCCCTGCAAAATGTGACCTTCCCTTACTGGCAGGTGAGCAGTTTCCGAAGTGAACAGA 5196
DB |||||
DB 4948 TCGTCTGCAAAATGTTACGTTTCCAACTGGAGTGTGCACAAACCTGCATGGGACATGGA 5007
QY |||||
QY 5197 CAGTTCACTGATGTCAGGTAGAGACCTCCAGC---AAGGATGATGTACCATTAACAC 5253
DB |||||
DB 5008 CAAGCTATCAAGTGGACATTTGAGTCTATAGCCGAGAAGAAGATGTGACATCACTTC 5067
QY |||||
QY 5254 CAACCTGTCTGCTCAGGCGCTTCTACACACTGCTGCAAGCAACCAAGTCCCGGAACAGCTA 5313
DB |||||
DB 5068 AATCTGTCTCGATCGATTCTTCTACACCATGGTTCAAGATCAGTTAAGAAACAGCTA 5127
QY |||||
QY 5314 CTACATCGGGCCGATGCTCTTGGCGCTGCTGCGCCAAACGGCATGGAGGTGGCGCT 5373
DB |||||
DB 5128 CCAGATTGTTATGACGCGCTCCCTCAGAAATTTATACGCCAGTGGCCTGCACTACACTA 5187
QY |||||
QY 5374 GCAGACTCAGCCCCACTTGTGGCTGGCACCGTCAACCCACCGTGGGCAAGAGGAATGT 5433
DB |||||
DB 5188 CCAACAGAGCGGACGTTCTGGCTGGCACCGCTAATCCGACGGTTGCCAAAGAAACAT 5247
QY |||||
QY 5434 CACGCTGCCCATCGCAACAGGCTCTCAACTGGTGGAGTGGCGCCGACGCGCAAGAGCAGG 5493
DB |||||
DB 5248 GACTTTGCTGGCAGAAACGGTCAAACTTGGTGAATGGAGATTCGCAAAAGAGCAAG 5307
QY |||||
QY 5494 TCGGGCCAGGTCACTGCTTTGGCGCGCTGGGGTGCACAAACGAATCTCTCTATC 5553
DB |||||
DB 5308 CCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGTTAATGGCAGAAACCTCTCTTC 5367
QY |||||
QY 5554 TCTGACTTTGATCGGCTAAACACGACAGAGAAGATCTATGATGACCAACCGCAAGTTTCA 5613
DB |||||

Db 5368 AGTTGACCTTTGATCGAACAAACAGACAGAAAAGATCTATGACGACCACCGTAAATTTCT 5427
Qy 5614 CTTTCGGGATTTCTGACGACAGCGGGCGGCCAGGCTCTGTCACCCAGCAGCAGGCT 5673
Db 5428 ACTGAGGATCCGCTACGACAGCTGCGGCACCGGACTCTCTGGCTGCCAAGCAGCAGCT 5487
Qy 5674 GAATGGTGTCAAGCTGACATACCTCCCTGGGGGTTTACATTCGTGGCATCCAGAGGGCAT 5733
Db 5488 GATGGCGGTCAATGTGACCTATTATCCACAGGTCAAAATGGCCAGCATCCAGCGAGCAC 5547
Qy 5734 CATGTCTGAAGAAGATGGAAATACGACAGCGGGCGGCAGATCCACAGGATCTTCGTGTA 5793
Db 5548 NACTAGCGAGAAGTAGATTATGACGACAGCGGAGGAGATCGTCTCGGGTCTTTGCTGA 5607
Qy 5794 TGGGAAGACATGGAGCTACACATCTTAGAGAAAGTCCATGGTGTGCTACTACACAGCCA 5853
Db 5608 TGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGGTCTCTGCTCTATAGCCA 5667
Qy 5854 GAGGACGATATCTTTGAGTTTCGAAGAATGACCGGCTCTCTCTCTGTGACGATGCCAA 5913
Db 5668 GCGGCGGTACATCTTCGAATACGATATGTGGGACCGGCTGTCTGCCATCACCATGCCAG 5727
Qy 5914 CGTGGCGGGCAGACACTAGAGACCATCCGCTCAGTGGGTACTACAGAAACATCTATCA 5973
Db 5728 TGTGGCTCGCACACCATCGCAGACCATCCGATCCATGGCTACTACCGCAACATATCAA 5787
Qy 5974 GCCCCTCAGGCGCAATGCTCAGTCATACAGGACTTCACTGAGGATGGGCACCTCTTCA 6033
Db 5788 CCCCCGGAAGCAACGCTCCATCATCACGACTACAAAGAGGGGCTGCTTACA 5847
Qy 6034 CACCTTCTACCTGGGCACTGGCGGAGGGTGATATACAGTATGGCAAACTGTCAAAGCT 6093
Db 5848 AACAGCTTCTTGGGTACAAGTCGAGGGTCTTATTCAAATACAGAAGGCAGACTAGGCT 5907
Qy 6094 GCGAGAGCGCTCTATGACACACCAAGGTCAAGTTTCACTATCACAGAGCGGAGCAT 6153
Db 5908 CTCAGAAATTTATATGATAGCAACAGTCAAGTTTACCTATGATGAAACAGCAGAGT 5967
Qy 6154 GCTGAAGACCATCAACCTACAGAAATGAGGGTTCACCTGACCATCCGCTACCGTCAGAT 6213
Db 5968 CTTAAAGACAGTAAACCTCCAGAGTGATGGTTTTATTGTCACCATTAGATACAGGCCAAAT 6027
Qy 6214 TGGGCCCCGTATGACACAGATCTCCGCTTCACTGAGGAAGGCATGTCACGCGCG 6273
Db 6028 TGGTCCCCGTATGACAGCGAGATTTCCGCTTTAGTGAAGATGGGATGTAATGCAAG 6087
Qy 6274 TTTTGACTACAACTATGACAAACAGCTTCCGGGTGACACAGCATGCAGGCTGTGATCAACGA 6333
Db 6088 ATTTGACTATAGCTATGACACAGCTTTTCGAGTGACCCAGCATGCAGGGTGTGATCAATGA 6147
Qy 6334 GACCCCACTGCCATTTGATCTCTATCGCTATGATGTGTCAGGCAAGACAGAGATTT 6393
Db 6148 AACGCCACTGGCTATTTGATCTGTATCAGTTTGTATGACATTTCTGGCAAAAGTTGAGCAGTT 6207
Qy 6394 TGGGAAGTTGGTGTCAATTTACTATGACATTAACAGATCATCACCAGCTGTCAATGAC 6453
Db 6208 TGGAAAGTTTGAGTTATATATATGATATTAACAGATCAITTTCTACAGCTGTAAATGAC 6267
Qy 6454 CCACACCAAGCATTTTGTATGATATGCGAGATGGAAGAGTGCAGTATGAGATCTTCGG 6513
Db 6268 CTATACGAAGCACTTTGATGTCTATGGCCGTATCAAGGAGATTCAAATATGAGATTTACG 6327
Qy 6514 CTCGCTCATGTACTGATGACCGTCCAGTATGATTAACATGGGGCGAGTATGAGAAGGA 6573
Db 6328 GTCGCTCATGTACTGATTAACAATTCAGTATGATAACATGGGTGGGTAAACCAAGAGAGA 6387
Qy 6574 GCTGAAGGTAGGACCCCTACGGCAATACCACCTCGGTACTCTCTATGAGTATGATGTGACGG 6633
Db 6388 GATTAAATAGGGCCCTTTGCCAACACCAACCAATATGCTTTATGATATGATGTGATGG 6447
Qy 6634 CCAGCTGCAGACAGTCTCCATCAATGACAGCCACTCTGGCGCTACAGCTACGACCTCAA 6693
Db 6448 ACAGCTCCAAACAGTTTACCTCAATGAAAAGATATGTGGCGGTACAACTACGATCTGAA 6507

Qy 6694 TGGGACCTTGCACTTACTTACTGAGCCCTGGGAAACAGTGCAGGCTCACACACTACGGTATGA 6753
Db 6508 TGGAAACCTTCCATTTACTGAAACCAAGTAACAGTGGCGCTCTGACACCCCTTCGCTATGA 6567
Qy 6754 CATCCGCGACCCGATCACTCGGCTGGGTGACGTGCAATACAAAGATGGATGAGGATGGCTT 6813
Db 6568 CCTGGGAGACAGAACTCACTCGACTGGTGTATCTCAATATCGTTGGATGAAGATGGTTT 6627
Qy 6814 CTTGAGGACAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGCTGCCCTCATCAAGC 6873
Db 6628 CTTACGTTCAAGGGCGACCGAAATCTTTGAAATATAGCTTCAAGGGGCTTCTAACTGAGT 6687
Qy 6874 CTACAAACCGGCTGGCAGCTGGAGTGTCAAGTACCGCTACGATGGCTGGGGCGGCGCT 6933
Db 6688 TTACAGTAAAGCAGTGGCTGGACAGTGTCTACCTTTATGACGGCTGGGAAGCGTGT 6747
Qy 6934 GTCCAGCAAGAGCAGCCACAGCCACACTGAGTCTTCTATGACAGCTTCAACCAACCC 6993
Db 6748 TTCTAGCAAAACAGTCTAGGACAGCAGCTGAGTTTTTTATGCTGACTTAACTTATCC 6807
Qy 6994 CACCAAGGTACCCACTGTCAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGA 7053
Db 6808 CACTAGGATTTACTGCTCTCAACCACTTCGAGTTCAGAAATCTCCCTGTATTAGA 6867
Qy 7054 CTTGCAAGGACACTCTTTTCCCATGGAGCTGAGCAGTGGTGTGATGATTTTACATAGCTTG 7113
Db 6868 TCTCCAAAGGACATCTTTTTCCTATGCAATCAGCAGTGGGATGAATCTATATTGCATC 6927
Qy 7114 TGAACACTCGGACCCCTCTTCTGCTTTAGTGGACAGGTTTGTAGATCAAGCAAT 7173
Db 6928 GGATAACACAGGACACCACTGGCTGTGTTTCAGTAGCAATGGGCTTATGCTGAAACAGAT 6987
Qy 7174 CTTGTACACAGGCTATGGGGAGATCTACATGGATACCAACCCCACTTTTCAGATCATCAT 7233
Db 6988 TCAGTACACTGCTATGGGGAATCTATTTGACTCTATATTGACTTTTCACTGTAAT 7047
Qy 7234 AGCTACCATGGTGGCTCTATGATCCACTCAACAGCTTGTCCACATGGGCGGCGAGA 7293
Db 7048 TGGATTTTCATGGTGGCTGTATGCCACTCAACAAATTAATCCACTTTTGGAGAAAGGA 7107
Qy 7294 TTATGATGTCTGGCGGACGCTGGAGTAGCCAGACACGAGCTGTGGAAGCACCTTAG 7353
Db 7108 TTATGACATTTTGGCAGGACGCTGGACACACCTGACATAGAAATCTGGA--AAAGAT 7164
Qy 7354 TAGCAGCAACGTATGCTTTTAACTCTATATGTTTCAAAAAACAAACCCCATCAGCAA 7413
Db 7165 TGGGAAAGGACCCAGCTCTTTTAACTTTGATGTTTGGAAATTAACACCTCTGCAAGCAA 7224
Qy 7414 CTCCAGGACATCAAGTGGCTTCATGACAGATGTTAACAGCTGGCTGCTCACCTTTGATTT 7473
Db 7225 AATCCATGACGTGAAAGATTAACATCAGATGTTAACAGCTGGCTGGTGCATTTGGTTT 7284
Qy 7474 CCAGCTACACACAGTATCCCTGGTTTATCCAAACACAGACATGGATGCCATGGAACCTC 7533
Db 7285 CCATCTGCACATGCTATTTCTGGATTCTCTGTTCCAAATTTGATTTAACAGACCTTC 7344
Qy 7534 CTACAGAGCTATCCACACACAGATGAAAACGAGGATGGGACACAGCAAGTCTATCCT 7593
Db 7345 TTACGA-----ACTTGTGAAGAGTCAGCAGTGGGATGATATACGCCCCATCTT 7392
Qy 7594 CGGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCTTTGTCCCTTAGAACGCTTTGA 7653
Db 7393 CGAGTCCAGCAGCAATGGCGGCGGCGCAAGGCTTCTCTGCTCGCTGGGAGATGCG 7452
Qy 7654 CCAGCTCTATGGCTCCAAATCACAGCTGCCAGCAGGCTCCAAAGACCAAGATTTGCG 7713
Db 7453 CGAGGTGCAGTGCAGCG 7509
Qy 7714 ATCCAGGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGGCTTGAAGGATGGCGGCTGAC 7773
Db 7510 CACGGTCAAGTCTGCTGATCGGCAAGGGCGCTCATGCTGGCGCTCAGCCAGCGCGCGCTGCA 7569

QY 7774 CACAGACATCATAGTGGGCAATGAGTGGGAGAGGTTGTCGCTTGTGACCA 7833
 Db 7570 GACCAACGTCTCAATCGGCAACGAGACTGATCAAGTGGCGCGTGTCTCAACA 7629
 QY 7834 TGCCCACTACCTAGAGAACCTGCACTTCCACATTCATGGGTGGATACCCATTACTTTGT 7893
 Db 7630 CGCCTTCTACCTGGAGAACCTGCACTTCCACATGGAGGCAAGCACGCACTACTTCAT 7689
 QY 7894 GAACACGAGACTTCAGAGGTGACCTGGCCATCTCTGGGCTCAGTGGGGGGGGCGAAC 7953
 Db 7690 CAAGACCAACCGCCGAGAGCGACCTGGGACACCTGCGGTGACCGGCGCGCAAGGC 7749
 QY 7954 CCGTGAAGTGGGTCAACGTCACTGTGTCACAGATCAACACAGTACTTAATGCGAGAC 8013
 Db 7750 GCTGAGAACGCGATCAACGTGACGGTGTGCGAGTCCACACGCTGGTGAACGCGAGGAC 7809
 QY 8014 TAGCGCTACAGACATCCAGCTCCAGTACGGGGCACTGTGTTGAACACACGCTACGG 8073
 Db 7810 GCGCAGGTTCCGCGACGCTGGAGATGCACTTCCGGGCGCTGCGCTGACCGTGCCTACGG 7869
 QY 8074 GACAACGTTGATGAGGAGAGGACGCGTCTCTGAGCTGCGCCGCGAGAGCGGTGG 8133
 Db 7870 CATGACCTTGACGAGGAGAGGCGCGATCTCTGAGCAGCGCGCGGCTGCG 7929
 QY 8134 CCAAGCTGGCCCGCGAGCAGCAGACTGCGGGAAGGGAGGAGGCTGCGGGCGTG 8193
 Db 7930 CCGGCGCTGGCGCGGAGCAGCAGCGCTGCGGCGAGCGGAGGCGCGCGCTCTG 7989
 QY 8194 GACGAGGGGAGAGCAGCAGGCTGCTGAGCAGCAGGCGCGGTCAAGCTACGACGCTT 8253
 Db 7990 GACGAGGGGAGAGGCGCGAGCTGCTGAGCAGCAGGCGCGGTCAAGCTACGACGCTT 8049
 QY 8254 TTTCGTGATCTCTGTCAGCAGTACCCAGAACTGTCAGACAGCGCCCAACATCCACTT 8313
 Db 8050 CTAGCTACTCTCGTGGAGCAGTACCCGAGCTGGCGAGCGCCCAACATCCAGTT 8109
 QY 8314 CATGACAGCAGCAGGATGGCCGAGGTGAC 8345
 Db 8110 CCGCGCAGAGCAGATCGCAGGAGGTAAC 8141

RESULT 8
 AAS14085
 ID AAS14085 standard; DNA; 9826 BP.

AC AAS14085;

DT 18-DEC-2001 (first entry)

DE Human FCTR3b DNA sequence.

XX Human; FCTR3b; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
 KW astrocytoma; congenital neonatal alloimmune thrombocytopenia; infection;
 KW neurological disorder; neurodegenerative disorders; nerve trauma;
 KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
 KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
 KW mental health condition; immunological disorder; allergy; infertility;
 KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
 KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
 KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
 KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
 KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
 KW Corneal dystrophy-Greenuw type I; Corneal dystrophy-lattice type I;
 KW Reis-Bucklers corneal dystrophy; cytosatic; immunosuppressive;
 KW antiasthmatic; antinfertility; antiinflammatory;
 KW antidiabetic; hepatotropic; virucide; ophthalmological;
 KW gynaecological; antinfertility; immunostimulant; auditory; haemostatic;
 KW gene therapy; FCTR3b; neurastin-like protein.

OS Homo sapiens.

XX Key Location/Qualifiers

PH 5'UTR 1. .279

FT CDS /*tag= b
 FT 280..8481
 FT /*tag= a
 FT /product= "Human FCTR3b"
 FT 8482..9826
 FT /*tag= c
 XX
 PN WO200166747-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001WO-US007160.
 XX
 PR 03-MAR-2000; 2000US-0186592P.
 PR 03-MAR-2000; 2000US-0186718P.
 PR 06-MAR-2000; 2000US-0187293P.
 PR 06-MAR-2000; 2000US-0187294P.
 PR 17-MAR-2000; 2000US-0190400P.
 PR 07-APR-2000; 2000US-0196018P.
 PR 03-JAN-2001; 2001US-0259548P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
 PI MacDougall J, Mishra V, Mezes PS, Rastelli L;
 XX
 DR WPI; 2001-596837/67.
 DR P-PSDB; AAU08580.
 XX
 PT Novel polypeptides designated as FCTR3 polypeptides, useful in detection,
 PT prevention and treatment of a broad range of pathological states.
 XX
 PS Claim 9; Page 33-35; 215pp; English.
 CC The invention relates to human FCTR3 polypeptides, FCTR1-FCTR7, and the
 CC nucleic acids encoding them. These sequences are useful for the treatment
 CC or prevention of numerous disorders including myelogenous leukaemia,
 CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
 CC alloimmune thrombocytopenia, neurological disorders, neurodegenerative
 CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
 CC Tooth neuropathy, demyelinating Gardner syndrome, familial
 CC myelodysplastic syndrome, mental health conditions, immunological
 CC disorders, allergy and infection, bronchial asthma, Avellino type
 CC eosinophilia, lung diseases, reproductive disorders, infertility, male
 CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
 CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
 CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
 CC infection, Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
 CC Corneal dystrophy-Greenuw type I, Corneal dystrophy-lattice type I and
 CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
 CC FCTR3b, a neurastin-like protein
 XX
 SQ Sequence 9826 BP; 2567 A; 2643 C; 2479 G; 2137 T; 0 U; 0 Other;

Query Match 40.5%; Score 3385; DB 5; Length 9826;
 Best Local Similarity 64.7%; Pred. No. 0;
 Matches 5393; Conservative 0; Mismatches 2790; Indels 154; Gaps 18;
 QY 35 ATGGACGTGAAGGAGAGAACCTTACCGCTCGTGTGACCGCGCGCGAGCGGCGG 94
 Db 280 ATGGATGTAAGGAGCGCGGACACCGCTCTTTGACAGAGGACGCTGTGGCAAGAGTGT 339
 QY 95 CGCTACACACCTCGTCCCGGACAGGAGGAGGCAAG---CCCCGCGAATTCCTAC 151
 Db 340 CGCTACACACCTCTCTCTGACAGTGAAGCTGCGGGTGGCCACACAGAAATCTTAC 399
 QY 152 AGCTCCAGGAGACCTTGAAGGCGCTACACAGGAGCGCGCGCTAGCTATGGCAGCGC 211
 Db 400 AGCTCCAGTGAGACTCTGAAGGCGCTATGACCATGACAGCAGGATGCTATGGAACCGA 459
 QY 212 GTCAAGGACATGTCCCGCAGGAGGCGGAGAAATCTGCCGACAGGTGCCAATTCACC 271
 Db 460 GTCAAGGACCTCATCCACCGGGAGTGATGAGTTTCTTAGCAAGGAACCACTTCACC 519

QY 272 CTGGGGAGCTGGGCTGGGAAGATTAACGCCCCCTCAGCGGACCCCTGTACCGAGACAG 331
 Db 520 CTTGCCGAACCTGGGCTCTGTGA---GCCCTCCCAACACCGAAGCGCTACTGTCCGAC 576
 QY 332 ATTGGCCTGCCCAATGGGCTACTTCAATGGGGGTGGCTCTGATGCCAGATGAGGCT 391
 Db 577 ATGGGATCCCTTACCAGGGCTACTCCCTTAGCAGAGGCTCTGAGCGGACTCGACACC 636
 QY 392 GACAGGCTGTGTCCCTTGAGACACCCCGTGTGTGGGGCCGGAGACACAGGCTCAGGG 451
 Db 637 GAGGAGGGATGTCTCCAGAACACCGCCATCAGATGTGGGGCAGAGGATAAATCCAGG 696
 QY 452 CGCAGCTCTCGCTGTCCAGCGGGCCAAATCCATCTCAACATCACCGACCGAGCAT 511
 Db 697 CGCAGTTCGGGCTGTCCAGTGTGGAACCTGGGCCCTTACCTGACTGTGACAAAC 756
 QY 512 GAAACACTGAGCTGATCATCCGGGGGCC-----TGAGAACACACCGGGCTCCGGAC 567
 Db 757 GAAACAAATCAGATGATGAGAACGGTGTCCCATTCACCTACATCCTGCGCTAGTCTC 816
 QY 568 GCCCGCGCGGCTCTCGACGCCCAACCCCAACACAGACACACCGGCTCCATTAA 627
 Db 817 CTCCTCATCTGTCTAGCTGCTTCCATTAATCTCCACAGTTAGTCTGCC-AGATGCC 875
 QY 628 CTCCTGAACCGGGCAACTTCAAGCCGAGGAGCAACCCAGCGCGGCCCCACCGACCA 687
 Db 876 ATTGCTAGACGACACCTCCCATCAATCATGACACACCAACCTGATGAGGATTTCTC 935
 QY 688 CTGCTCTCCGAGAGCCCTTGGCGGGCGCCAGAGCTCTGCCACCGCCAGGAGAA 747
 Db 936 CCCCATAATCATCTGCTCAGAGATGCTCAGGCGCCCGAGCAAGCTCCAGCAGTGGCC 995
 QY 748 CTGGCTGTCAACAGCAACATCCCTCGAGACAGGAAACCTAGGCAAGAGCATTCCT 807
 Db 996 TCCGAAACACACGCTGCTGAGGCTCTGAGGCC-----CTCTCCC 1037
 QY 808 AGGCAATTTGAGGACAACTCATTTAGATGACATTTCTGGCGCCCTCCGCCATGATGG 867
 Db 1038 ACCCCTCACAACACACAGCTGTCCCATCAGCACTCGTCCGCCAACTCCCTCAACAGAA 1097
 QY 868 GCTTACAGTACGGGCACTTCTCTTCAAGCTTGAGGACCTCCCGGCTCTTCTGAC 927
 Db 1098 CTCATGACCAATCGCGGAGTCTAGATCCAGCCCGGCCCGAGCGCCCAATGACCTGGC 1157
 QY 928 CACATCACAGGTTACCACTGACGTCCAGACAGTGTACTCTCTCGCCCGGACCCCT 987
 Db 1158 CACACACAGAGTCCGTTTCACTTCAGGACAGCTGGGTGCTAAACAGCACTGCTCACT 1217
 QY 988 GCCCGGAGACCTTTGCGCGGCGGCTTTAACTCAAGAGCCCTTCAAGTACTGTAA 1047
 Db 1218 GGAGACCGGCACTTCTCTTCAAGACCTCTCGGGAGCACACCT-----T 1265
 QY 1048 CTGGAAGTGGCAGCCCTGAGCGCATCGTCTATCTCAGCCACTCTGGTCACTCTGGC 1107
 Db 1266 GTTACAGAGCTCTTCCCGGATACCTTTGACCTCAGAACGGTTTACAGCCCGCCGCC 1325
 QY 1108 ATACTTTGGCCATGACCTGTTTGGCTAACTGGCACTGACGCGGATGAGGGGCA 1167
 Db 1326 CGCCTGTGCCAGGAATCTTTCTCCA-----GGAAGGCTTTCAA 1367
 QY 1168 GATGTATGATCAGGAGGACACAGCAGAGTGGCTGTGCCAACCGACGCTCC-C 1226
 Db 1368 GCTGAAGAGCCCTCCAATATCTGAGCTGGAATGTCTGCCCTCTCCGCCATTTGCCG 1427
 QY 1227 TATACCCCTCAGGGGGCACTGGCTTAGAGACCCCTGACAGGAAAGGCAAGGACACAG 1286
 Db 1428 GSCCTCTCTTGGCTATTTTGTGGCTATTTTATAGTGCCCTGGTGTGAAACAG 1487
 QY 1287 AAGGAAAGCCCATGATTTCTTTCCAGAGGACAGTTTCATAGATTTGAGGAAATGATG 1346
 Db 1488 CAG-----CATAGACAGTGGTGAAGCAGAA 1513

QY 1347 TGGGAAGGAGAGCTCCAGAGATTCCTCTCTGGCACTTTCTGGAGATCTCAAGTGTCA 1406
 Db 1514 TTGGTGGCGGGTAACACAGAGAGTCCACAGAGGTGTTTTGGAGGTCAAAATTCACA 1573
 QY 1407 TAGACCATCTGTGATCTGAAATCAATGTGTCTCTGGAAAGGAGCAGCCTCTCTTTGG 1466
 Db 1574 TCAGTCAGCCCCAGTCTTTAAAGTTCAACATCTCCCTCGGAAGGACGCTCTCTTTGG 1633
 QY 1467 TTTATGGGAGAAAGGCTCCCTCTTACATACACAGTTTACATTTTGGAGCTCTCTG 1526
 Db 1634 TTTACATAAGAGAGACTTCCACATCTCATGCCAGTATGACTTCATGGAACGCTCTG 1693
 QY 1527 ATGGCAGGAGGCTCTTAACCCAGGAGCGCGAGCTAGAGGGACCCCGCGCAGCTCTC 1586
 Db 1694 ACGGGA-----AGGAGAGTGGAGTGTGTTGAGTCTCCACAGGAACGCC 1738
 QY 1587 GGGGAACCTGTGCCCCCTCCAGCATGACAGAGCTTCACTCAGATTTTGGATTTCAAGAA 1646
 Db 1739 GGAGCATACAGACCTTGGTTTCAAGATGAAGCGCTGTTTGTGCACTACCTGAGTGTGG 1798
 QY 1647 TCTGGCACTTGGCTTTTACAATGACGGAAGAGTCAAGTGGTTTCTTTCTCAACA 1706
 Db 1799 TGTGGCATCTGGCTTCTACAATGATGGAAGAAACAAAGAGATGTTTCTTCAATCTG 1858
 QY 1707 CTGCCATTGAGTGGGTGATAACTGCCCAGCAACTGCTTATGCAATGATGATGATCT 1766
 Db 1859 TTGTCTAGATTCACTGTCAGGACTGTCCACGTAACCTGCAATGGAAATGATGATGTTGT 1918
 QY 1767 CTGGGACCTGCCACTCTTCTGGGTTTCTGGGCCCCGACTGTGGCAGAGCTCTCTGCC 1826
 Db 1919 CCGGGTGTGTCACTGTTTCCAGGATTTCTAGGAGCAGACTGTGTAAGCTGCTCTGCC 1978
 QY 1827 CCGTGTCTGTAGCGAAATGCCAATACATGAAGGAGAGATGCTTTGTGCCACAGTGGCT 1886
 Db 1979 CTGTCTGTGCACTGGGATGGACATATTTCAAAGGACGCTGCCAGTCTCAGCGGCT 2038
 QY 1887 GGAAGCGGCTGAGTGGATGTGCCCAACCAACAGTGTATGATGTGGCTCGAGCAACC 1946
 Db 2039 GGAAGGTGAGAGTGGAGCTGCCATGAATCAGTGTGATGATGCTCTCTCGGGGGCC 2098
 QY 1947 ATGGCACTGTGATCAGGCACTGTGATCTGCAACCTCGCTACAGGCGGAGAGCTGTG 2006
 Db 2099 ACGCTCTGTGATGAGGAACTGTGTCTGTCTGTGCTGCTACAAAGGAGCACTGTG 2158
 QY 2007 AGGAAGTGAATGATGAGACCCCACTATGTTTCAAGCGGGGTGTCTGCTGAGAGGCAAT 2066
 Db 2159 AGGAAGTGAATGATGAGACCCCACTATGTTTCAAGCGGGGTGTCTGTTGAATGAGAA 2218
 QY 2067 GCAATTTGCTTTGGGATGGGAGGACCAACTGCGAGACCCCGAGGCACTGCTTAG 2126
 Db 2219 GCTGTGACGCTGCTGGGTTGCTGAACTGTGAGTGGGAGGTTCCAGTGTGCCAG 2278
 QY 2127 ACCAGTGTTCAGGCAACGGAACCTTCTCCGAGACACCGGGCTTTGACGCTGTGACCAA 2186
 Db 2279 ACCAGTGTGAGTGGGATGGCATGCACTGCTGCTGACACGGGCTCTGCACTGCACTCCA 2338
 QY 2187 GCTGAGTGAACAGACTGTTTATGAGATCTGTCTGCGGACCTGTGTGTGGCCATGGG 2246
 Db 2339 ACTGGATGGGTCCGACTGCTCTGTTTGAAGTGTCTCAGTAGACTGTGCACTCAGCGG 2398
 QY 2247 TGTGGTGGGGGCACTCGCTGCGAGATGGCTGGATGGGGGAGCTGCGGAGCAGCTGCG 2306
 Db 2399 TCTGCAATGGGGAGGCTGCGCTGTGAAGAGGCTGGACAGGCGAGGCTGTGACGAGC 2458
 QY 2307 GGGCCTGCCACCGGCTGTGCGAGCATGGGACCTTGGCGGAGCGGCAAGTGGAGTGA 2366
 Db 2459 GCGTGTGCCACCCCGCTGCTTGTGAGCAGCGGACCTGTAAGATGGCAATGTGAATGCC 2518
 QY 2367 GCGCTGGCTGGATGGGAGACACTGACCATCGCTCACTATCTGGATAGGTTAGTAAAG 2426
 Db 2519 GAGAGGCTGGAATGTTGAACACTGACCATTTGGTAGGCAAAACGCGGAGCAGCCGAAC 2578
 QY 2427 AGGTTTGGCTGTGCAATGGCAACGGCAGATGTACTTTAGACCTTGAATGGTTGGC 2486

Db 2579 ATGGCTGCCCTGACCTGTGTGCAAGCGTACGGGAGATGCACACTGGTCAGAAACAGCTGGC 2638
 QY 2487 ACTGGCTGTGCCAGCTGGCTGGAGAGAGCTGCTGTGACACTTCCATGGAGACTGCTCT 2546
 Db 2639 AGTGTGTGTGGCAGACCGGCTGGAGAGGGCCGGATGCAAGCTTGGCCATGGAACTTCCCT 2698
 QY 2547 GCGGTGACAGCAAGACAATGATGAGATGCCCTGGTGGACTGTCATGGACCCCTGACTGCT 2606
 Db 2699 GTGTGTATAAACAAGATATAGGGAGATGAGCTGTGGTGGATGTTGGACCTGACTGCT 2758
 QY 2607 GCTCCAGCCCTGTGCCATATCAACCCGCTGTGCTGTGGCTCCCTTAAACCCCTGTGAACA 2666
 Db 2759 GCTTCAGTACGCTGTGACAGACAGCTGCTCTGCGGGGGTCCCGGAGCCCACTGGACA 2818
 QY 2667 TCATCCAGGAGACAGCTCCCTGTGTACAGAGAGAACCTTACACTCTCTTATGACCGCA 2726
 Db 2819 TCAITTCAG-----CAGGGCCAGAGGATTTGGCCCGCAGTGAAGTCTCTTATGACCGTA 2872
 QY 2727 TCAAGTTCCTGTGGGCGAGGACAGCAGCAATATAATCCCGGGGAGAACCCCTTTGATG 2786
 Db 2873 TCAAGCTCTTGGCAGGCAAGATAGCACCCACATCATCTCTGGAGAGAACCCCTTTCAACA 2932
 QY 2787 GAGGCGATCTGTGTATTCTGTGGCCAGATGATGATCAGATGGAGAACCCCTGTGTTG 2846
 Db 2933 CGAGCTTGTGTTCTCTCATCCGAGGCGCAAGTAGTAACCTACAGATGGAACCTCCCTGTGTCG 2992
 QY 2847 GTGTGAACATCAGTTTTCTCAATAACCTCTCTTTGGATATACAAATCAGCAGGCAAGATG 2906
 Db 2993 GTGTGAACGTGTCTTTGTCAAGTACCAAAATACGCTACACCATCACCCGCGCAGATG 3052
 QY 2907 GCAGCTTTGACTGTGTGCAAAATGGGCGCATCTCCATCATCTCTCGGTTTCAGCGGGCAC 2966
 Db 3053 GCAGCTTCGACTGATCGCAATGAGGTGTCTCTTCACTTACACTTTTGGAGCGAGCC 3112
 QY 2967 CTTCATCACACAGGAGCACACCTGTGGCTGCCATGGATCGGTTCTTTGTATGGAAA 3026
 Db 3113 GTTCATGAGCCAGGAGGCGCACTGTGTGGCTGCGTGGAGACAGCTTTTACGCCATGGACA 3172
 QY 3027 CCATCATCATGAGACATGAGGAGAAATGAGATTCGCCAGCTGTGACCTGAGCAAAATTTGCC 3086
 Db 3173 CCCTGGTGTAGAACACCGAGGAGAACTCCATCCCGCAGCTGTGACCTCAGTGGGCTTTGTCC 3232
 QY 3087 GCGCCAAACCCAGTCTCTCCATCCCACTGAGAGTCTTCCGCGAGCTCTCTGTGCGAGAGA 3146
 Db 3233 GGCCTGATCCAATCATCATCTCTCCCGCAGTCTCCACTCTTTAGTGTGCTGCCCTGGGC 3292
 QY 3147 AAGGCCCATTTGTCCGGAAATTCAGGCTTTTGCAGGAGGAAATCTCTATCTCTGGCTGCA 3206
 Db 3293 AGAATCCCATCTGCTGAGACCCAGGTTCTTCATGAAGAAATCGAGCTCCCTGGTTCCA 3352
 QY 3207 AGATGAGCTGAGCTACTGAGAGCGGAGCCCTGCTGACAAATCTGTCTGAGAGATCA 3266
 Db 3353 ATGTGAACCTTCTGATCTGAGCTCTAGAACTGAGAGGTTACAACTGCTGAGAGATCA 3412
 QY 3267 GCCTCACCCACCGACCAATCCCTTCAACCTCAAGAGTGCACCTCATGCTAGCGGTGG 3326
 Db 3413 CCATGACCCATCCACAGTGCCTTGAACCTCATATTAGGTTTCACTGATGGTGTGCTG 3472
 QY 3327 AGGCGGCTCTTTCAGGAAGTGGTTGCTGTGAGCGCCCGACACCTGTCTCTATTATTTCATTT 3386
 Db 3473 AGGGGCACTCTTCCAGAAATTCATCCAGGCTTCTCCCAACCTGGCCCTCCACCTTCATCT 3532
 QY 3387 GGGACAAGACAGAGCTCTACAAACAGAGGTGTTGGGCTTTTCAGAGGCTTTGTTTCCG 3446
 Db 3533 GGGACAAGACAGAGTGGTATGGCCAAAGGTGTATGACTCTGATGCTGTGTGTCTG 3592
 QY 3447 TGGGTTATGAATGAATCTGCGCCAGATCTAATCTGTGGGAAAAAGAACCAAGATGC 3506
 Db 3593 TCGGTTTGAATATGAGACCTGTCCAGTCTAATCTCTGGGAGAAAGACAGCCCTCC 3652
 QY 3507 TGCAGGCTATGAATGACGCGTCCAGCTTGGAGGATGGAGCTGAGCAAAACATCATG 3566

Db 3653 TTTCAGGATTCGAGCTGGACCCCTCAAACCTCGTGGTGGTCTCCCTAGACAAACACCACA 3712
 QY 3567 CCTCAACATTCAAAGTGGTATCTCTGCACAAAGGAATGGGGAGAACACAGTTGTGTCTC 3626
 Db 3713 TCTCAATGTTAAAGTGGATCTTACACAAAGGCACTGGGGAAACACAGTTCTTGACCC 3772
 QY 3627 AGCAGCTCTCTCATTTGGGAGCATCATGGGCAATGGCGCGGAGAGATCTCTCTGCC 3686
 Db 3773 AGCAGCTGCCATCATCACAGCATCATGGGCAATGGTCCGCGCGGAGCATTTCTGTCT 3832
 QY 3687 CCAGCTGCAACGGCCTTCTGTAGCGCAACAGCTCTCTGGCCCGCAGTGGCCCTCAGCTGTG 3746
 Db 3833 CCAGCTGCAACGGCCTTCTGTAGCGCAACAGCTCTCTGGCCCGCAGTGGCTCTGTGTGTG 3892
 QY 3747 GCTCTGACGGGAGCTCTATGTGGGTGATTTCAACTACATTAGAGGATCTTCCCTCTCTG 3806
 Db 3893 GAATCGATGGGAGCTCTATGTGGGTGATTTCAATTACATCCGAGCATCTTTCCCTCTC 3952
 QY 3807 GAAATGTCAACCAATCTCTAGAGCTGAGAAATAAGATTTTCAGACATAGTCACAGTCCAG 3866
 Db 3953 GAAATGTCAACCAATCTCTAGAGCTGAGAAATAAGATTTTCAGACATAGTCACCAACCCAG 4012
 QY 3867 CACACAAATACTACTCTGCGCCACAGACCCCATGAGTGGGCGCTCTCTCTCTGACAGCA 3926
 Db 4013 CACACAAATACTACTCTGCGCCACAGTGGACCCCGTCTCGGCTCTCTCTCTGACAGCA 4072
 QY 3927 ACAGCGCGGCTCTTTAAATCAAGTCCACTGTGTGGTGGTGAAGGACCTTGTCAAGAACT 3986
 Db 4073 ACAGCAGGAGAACTTACCGCGTCAAGTCTCTGAGTGGAAACCAAGACCTGTGCGAAT 4132
 QY 3987 CTGAGGTGGTTCGGGGAGAGTGCAGTGGCTCCCTTTTGTATGACACTCGCTCGCGGG 4046
 Db 4133 CGGAAATTTGTGCGAGGAGCGGAGAGAGTGTCTACCTTTGTATGAAGCCGCTGCGGG 4192
 QY 4047 ATGGTGGGAGGCGCACAGAACCACTCACCAATCCCGAGGGTATTACAGTGCAGCAAGT 4106
 Db 4193 ATGGAGGAGAGGCCATAGATGCAACCTGATGAGCCCGAGAGGTATTGCGATGACAAGA 4252
 QY 4107 TTGGGCTGATCTATCTTCTGATGGCACCATGATCAGACGATCGATCAGAAATGGATCA 4166
 Db 4253 ATGGGCTCATGTACTTGTCTGATGCGCACCATGATCGGAAAGTTGACAGAAATGGAATCA 4312
 QY 4167 TCTCACACCTGCTCGCTCTAATGATCTACATCAGCCCGCCACTCAGCTGATGATCTG 4226
 Db 4313 TCTCACACCTGCTCGCTCTAATGATCTAATGACCTCAGCTCGGCGCGCTGAGCTGTGATCCA 4372
 QY 4227 TCATGGATATTTCCAGGTAAAGCTGAGTGGCCCGCCACAGACTTAGCCATCAACCCAAATGG 4286
 Db 4373 GCATGGATGTAGCCCGAGGTTCTGTGGAGTGGCCAAACAGACCTTGTCTCAATCCCATGG 4432
 QY 4287 ACAACTCATTTATGCTCTCGACAACTGTGGTCTCTGCAATCTCTGAAAACCCAGG 4346
 Db 4433 ATAACTCTCTGTATGTTCTAGAGAACTATGCTCTTTCGAAATCACCGAGAACCCCAAG 4492
 QY 4347 TCGCATTTGTCTCGCGGAGGCGCCATGCACTGCCAGTCCCTGGGATTCACCACTTCTCTG 4406
 Db 4493 TCAGCATATTTCGGGAGCGCCCATGCACTGCCAAGTCTCTGGCATTTGACTACT--CAC 4549
 QY 4407 TAAAGCAAGTGGCCATCCACGCAACCTGGAGTGCAGCCACCGCTTTGGCTGTCTTTCACACA 4466
 Db 4550 TCAGCAAACTAGCAATTCACCTCTGCGCTGGAGTGCAGCGAGTGCCTTGGCAATTTCTCACA 4609
 QY 4467 ATGGGCTCTGTATTTGCTGAGACTGTATGAGAAAGATCAACCGCATCAGGAGGTCGA 4526
 Db 4610 CTGGGCTCTCTACATCACTGAGACAGATGAGAGAGATTAACCGTCTACGCCAGGTAA 4669
 QY 4527 CCATGATGAGAGATCTCACTCTGCTGGGGCCCCCAGTGGCTGTGACTGTGTAATAATG 4586
 Db 4670 CAACCAACGGGAGATCTGCTCTTTAGCTGGGGAGGCTTCGAGCTGGCAAAACG 4729
 QY 4587 ATGCCAACTGTGATTTTCTGAGAGCATGTTTATGCCAAGGATGCAAGATTTAAATA 4646
 Db 4730 ATGTCAATTTGCACTGTCTATTTCAGAGATGATGCTCAGGACTGATGCCATCTTGATTT 4789

QY	4647	CCCCATCTTCCTTGGCTGTGTCTGATGGGAGCTCTAGTGGCCGACCTTGGGAACA	4706	QY	5724	AGAGGGGCATCATGTCTGAAAGAAATGGAATACGACGAGCGGGCCGATACATCAGGA	5783
DB	4790	CCCCATCACTTTAGCTGTAGCTCCAGATGTACCATTTATTCAGATCTTGGAAATA	4849	DB	5870	ACGCTGGGCCCATGAGCGAGAGAGACAGACATCGAACAGCGCCGATCGTGTCCCGCA	5929
QY	4707	TCCGAATTCGGTTTATCCGGAAGAACAGCCCTTCTCAACACCCAGAACATGATGAGC	4766	QY	5784	TCTTGGCTGATGGGAAGACATCGAGCTACACATCTTAGAAGTCCATGGTCTGCTTAC	5843
DB	4850	TTCCGATCAGGGCGGTGAGCAAGAACAGCCCTTCTTAATGCTTCAACCCAGTATGAGG	4909	DB	5930	TGTTGGCTACGGGAAGTGTGGAGCTACTCTACTCTTGACAAGTCCATGGTCTCTCTGC	5989
QY	4767	TGCTTACCAATTCACAGAGCTCTATCTGTTTGATACACCGGCAAGCCTGTACA	4826	QY	5844	TACACAGCCAGAGGAGTATATCTTTAGTTTCGACAGAAATGACCGCTCTCTTTGTGA	5903
DB	4910	CTGATCCCGGAGAGCAGAGTTATATGTTTCAACGCTGATGGCATCCACCAATACA	4969	DB	5990	TTGAGAGCCAAAGTCAGTATATATTTAGTATGATCTCTGACCGCTCTCTTGGCGTCA	6049
QY	4827	CCCAAGCCTCCACAGAGACTACCTGTACAACTTCACTACCTAGGAGCGGACCA	4886	QY	5904	CGATGCCAAAGCTGGCGGGGACACATAGAGACCATCCGCTCAGTGGGCTACTACAGAA	5963
DB	4970	CTGTGAGCCTGGTGACAGGGAGTCTTGTACAAATTTACACATATGACTGACAATGATG	5029	DB	6050	CCATGCCACGCTGGCGGGACAGCATGTCCACACACACCTCCATCGGCTACATCCGTA	6109
QY	4887	TCACACTCATACAGACAAATGGAACATGTTAAATGTCGCGGAGACTCTACTGGGA	4946	QY	5964	AGATCTATCAGCCCTGAGGCAATGCTCAGTATACAGACATTCCTAGGATGGGC	6023
DB	5030	TCACCTGAATGATGACAAATGGAATTCCTTGAAGATCCGTCGGGACAGCAGTGCCA	5089	DB	6110	ATATTTCAACCCGCTGAAAGCATGCTTTCGGTCTATCTTGACTACAGTATGACGGCC	6169
QY	4947	TGCCCTCTGGCTGGTCCAGATGGCCAGGTGTAAGTGGTGACCATGGGACCAACA	5006	QY	6024	ACCTCCTTCAACCTTCTACCTGGGCACTGGCGGAGGTGATATACAAATGATGGCAAC	6083
DB	5090	TGCCCCCTCACTGCTCATGCTGACACACAGATCATCACCTCACCGTGGGACCAATG	5149	DB	6170	GCATCTGAAGACCTCTCTTTTGGGCAACGAGCGGAGGTCTTCTACAAATGATGGAAAC	6229
QY	5007	GTGCACTCAAGAGTGTACCAACACAGGACACGAGTTGGCCCATGATGACATACCATGGCA	5066	QY	6084	TGTCAAAGCTGGCAGAGACGCTCTATGACACACCAAGGTCTAGTTTCACTATGACGAGA	6143
DB	5150	GAGGCCTCAAAAGTGTGTCCACACAGAACCTGGAGCTTGTCTCATGACCTATGATGGCA	5209	DB	6230	TCTCCAAAGTATCAGAGATTGTACGACAGTACCGCCGCTACCTTCGGGTATGACGAGA	6289
QY	5067	ATTCGGGCTTCTGCAACCAAGCAATGAACGGATGGACAAACATTTTATGAGTACG	5126	QY	6144	CGGACGGCATGTGTAAGACCATCAACCTACAGAAAGAGGCTTCACTCCACCATCCGCT	6203
DB	5210	ACACTGGGCTCTGCGCCACCAAGAGCGATGAACAGGATGGACGACTTCTTATGACTATG	5269	DB	6290	CCACTGGTGTCTTGAAGATGGTCAACCTCCAAAGTGGGGCTTCTCTCCACCATCAGGT	6349
QY	5127	ACAGCTTTGGCGGCTGACAAATGTGACTTCCCTTACTGGCCAGGTGAGCAGTTTCCGAA	5186	QY	6204	ACCGTCAGATTGGGCCCCCTGATTACCGACAGATCTTCCGCTTCACTGAGGAAGGATGG	6263
DB	5270	ACCACGAAGCGCTGACCAAGCTGACGCGCCCCACGGGGTGGTAACAGTCTGCACC	5329	DB	6350	ACCGGAAGATTGGCCCCCTGGTGGACAAGCAGATCTACAGGTTCTCCGAGGAAGGATGG	6409
QY	5187	GTGATACAGACAGTTCACTGATGTCTCAGTGTAGAGACCTCCA---GCAAGATGATGTCA	5243	QY	6264	TCAACCCCGTTTGTACTACAACCT---ATGACAAAGCTTCCGGGTGACAGCATGCAGG	6320
DB	5330	GGGAATGGAGAAATCTATTACCAATGACATTTGAGAACTCAACCCGATGATACGCTCA	5389	DB	6410	TCAAATGCCAGTTTGTACTACCTATCATGACAAAGCTTCCGATCGCAAGCATCAAGC	6469
QY	5244	CCATAACCAACCACTGTGCTCTCAGCGGCTTCTACACACTGTGCAAGACCAAGTCC	5303	QY	6321	CTGTGATCAACAGAGACCCCACTGCCATTTGATCTCTATCGCTATGATGATGTGTAGGCA	6380
DB	5390	CTGTATCATCAACCACTCTCTTCACTAGAGGCTCTTACACAGTGGTACAAATCAAGTTC	5449	DB	6470	CCGTCTAAGTGAGACTCCCTCCCGTTGACCTCTACCGCTATGATGATGATTTCTGGCA	6529
QY	5304	GGAAAGCTATACATTCGGGGCGGATGGCTCTTGGCGCTGCTGCTGGCCCAACGGCATGG	5363	QY	6381	AGACAGAGAGTTTGGGAAGTTTGGTGTCTATTCTATCATCATTAACCATGATCATCACCA	6440
DB	5450	GGAAAGCTATACAGCTCTGTAAATATGTTACCTTGGGGTGTATGTCTTAATGGGATGG	5509	DB	6530	AGGTGAACACACTTTGTGAAGTTTGGAGTCACTTATTATGACATCAACCAAGATCATCACC	6589
QY	5364	AGGTGGCGCTGCAGACTGAGCCCCACTTGGCTGGCTGGCACCGTCAACCCACCGTGGCA	5423	QY	6441	CAGCTGTCTATGACCCACACCAAGCATTTTGTGATGATGTCATGCGAGGATGAAGGAAGTCCAGT	6500
DB	5510	GTATCAGCTTCCACAGGAGCCCCATGCTTCTAGCGGGACCATCAACCCCACTTGGAC	5569	DB	6590	CTGCCGTGATGACCTTCAGCAACACTTCGACACCCCATGGCGGATCAAGGAGGTCCAGT	6649
QY	5424	AGAGGAATGTACGCTGCCCCATCGAACACGGGCTCAACCTGGTGGAGTGGCGCCAGCGCA	5483	QY	6501	ATGAGATCTTCCGCTCGCTCATGTACTGGATGACCTCCAGTATGATTAACATGAGGCGGAG	6560
DB	5570	GCTGCAACATCTCCCTGCTATGGAAGTGGCTTAACTCCATGAGTGGCGCTTAAGAA	5629	DB	6650	ATGAGATGTTTCCGCTCCCTCATGTACTGGATGACGCTGCAATATGACATGAGGAGGG	6709
QY	5484	RAGACAGGCTCGGGCCAGTCACTGTCTTTGGCGCGGCTGCGGGTGCACAAACCGAA	5543	QY	6561	TAGTGAAGAGGAGCTGAGGTAGGACCTTACCCCAATACCACTCGCTACTCTCTATGAGT	6620
DB	5630	AGGAACAGATTAAAGGCAAGTCAACATCTTTGGCAGGAAGCTCCGGTCCATGGAAGAA	5689	DB	6710	TGATCAGAGGAGCTTAAACTGGGGCCCTATGCCAATACCAAGAGTACACCTATGACT	6769
QY	5544	ATCTCTATCTCTGACATTTGATCGGTAACACGACACAGAGAAGATCTATGATGACCAAC	5603	QY	6621	ATGATGCTAGCGCCAGCTGCAGACAGTCTTCATTAATGACAAAGCCACTCTGGCGCTACA	6680
DB	5690	ATCTCTTGTCCATTTGACTATGATCGAAATATTCGACTGAAAGATCTATGATGACCAAC	5749	DB	6770	ACGATGGGACGGGAGCTCCAGAGGTGGCGCTCAATGACCCCGGACCTGGCGGTACA	6829
QY	5604	GCAAGTTACCTCTCGAATTTGTACGACCGAGCGGGCGGCGCCAGCTCTGGTCAACCCA	5663	QY	6681	GCTTACGACCTCAATGGGAACCTGACTTACTGAGCCCTGGGAACAGTGCACGCTCACAC	6740
DB	5750	GGAAGTTACCTCTGAGGATCAATTTATGACAGGTGGCGCGCCCTCTCTGGTGGCCA	5809	DB	6830	GCTATGACCTTAATGGAAATCTCCATTTACTGAACCCAGCAAGTGTGGCGCTCATGC	6889
QY	5664	GCAGAGGCTGAATGCTGTCAACGTCATCTCCCTGGGGTTTACATTTGTCGATCC	5723	QY	6741	CACCTACGCTATGATCCCGGACCGCATCACTCGGTGGGTGACGTGCAATACAAAGTGG	6800
DB	5810	GCAGGGGCTGCGACGCTGTCAACGCTGTACATCTTTCAATGGGCGCCTGGCTGGCTTC	5869	DB	6890	CTTGGCTATGACCTCCGGGATCGGATTAACAGACTCGGGGATGTGCACTACAAATTC	6949
				QY	6801	ATGAGGATGCTTCTTGAGCAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCC	6860

Db 6950 ACAGCATGGCTATCTGTCGACAGAGAGGGTCTGACATCTTCCATATACATTTCCAAAGGCC 7009
 QY 6861 TGCTCATCAAGGCTTACAAACGGGGTGGCAGCTGGAGTGTCAAGTACCGTACGATGGCC 6920
 Db 7010 TCCCTAACAGAGCCTTACAAAGGCGAGCGGTGGAGTGTCCAGTACCGCTATGATGGCG 7069
 QY 6921 TGGGGGGGGCTGTCCAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 6980
 Db 7070 TAGGACGGGGCTTCTTACAGACCAACCTGGGCCACCACTGGAGTACTTCTACTCTG 7129
 QY 6981 ACTTGACCAACCCCAACCAAGGTCAACCACTGTACAACTTCCAGCTCTGAGATCACCT 7040
 Db 7130 ACCTCCACACCGAGCGCATCAACCATGCTACATCACTCCAACTCGAGATTACCT 7189
 QY 7041 CCCTCTACTAGACTTGAAGACACCTCTTTGGCATGGAGCTGAGCAGTGGTATGAGT 7100
 Db 7190 CACTGTACTAGACCTCCAGGGCCACCTCTTTGGCATGGAGCAGCAGTGGGAGGAGT 7249
 QY 7101 TTTACATAGCTTGTGACAACTCGGACCCCTCTTGTCTCTTTAGTGGACAGGTTTGA 7160
 Db 7250 ACTATGTTGCCCTGATACACAGGACTCTCTGGCTGTGTTGAGCATCAACGGCTCA 7309
 QY 7161 TGATCAAGCAATCTGTACACAGCCTATGGGAGATCTACATGATACCAACCCCACT 7220
 Db 7310 TGATCAACACAGCTGCAGTACACGGCCTATGGGAGATTTATATGACTCCAAACCCGACT 7369
 QY 7221 TTCAGATCATATAGCTTACCATGGTGGCTCTATGATCCACTACCAAGCTGTCCACA 7280
 Db 7370 TCCAGATGGTCAATGGCTTCCATGGGACTCTATGACCCCTGACCAGCTGGTCCACT 7429
 QY 7281 TGGCCGGCGAGATTATGATGTGTGGCGGACCTGTGACTAGCCAGACCAAGCTGT 7340
 Db 7430 TCACCTCAGCGTGAATATGATGTGTGGCAGAGCATGAGACCTCCAGACTATACCATGT 7489
 QY 7341 GGAAGCCTTTAGTAGCAGACAGCTATGCTTTTAACTCTATATGTTTCAAAACACA 7400
 Db 7490 GGAAGAACGT---GGGCAAGAGAGCGGCGCCCTTTAACTGTATATGTTTCAAGAGCAACA 7546
 QY 7401 ACCCATCAGCAACTCCAGGACATCAAGTCTCTCATGACAGATGTTTAAAGCTGGCTGC 7460
 Db 7547 ATCTCTCAGCAGTCACTAGATTGAGAACCTAGCTGACAGATGTGAAGCTGGCTTG 7606
 QY 7461 TCACCTTTGGATTCAGCTACACAGGTGATCTCTGTTATCCAAACCAAGCATGATG 7520
 Db 7607 TGATGTTTGGATTTCAAGCTTACCAACATCACTCTGCTTCCCGAGAGCCTCAAAATGATT 7666
 QY 7521 CCATGGAACTCTCTACAGCTATCCACACACAGATGAAACCGCAGAGTGGGACACA 7580
 Db 7667 TCGTGGCTCTCCCTATGA-----ATTGTGAGAGATCAAGCAAGTGAATG 7714
 QY 7581 GCAAGTCTATCTCGGGTACAGTGTGAAGTACAGAGCAGCTCAAGGCCCTTTGTCACT 7640
 Db 7715 GACAGCTCATTTACAGGTGTCCACAGACCAACAGAGAGACATAACAGGCCCTTCATGCTC 7774
 QY 7641 TAGAACGTTTGAACAGCTCTATGGCTCCCAATCAACAGTCCAGCAGCAGCTCCAAAGA 7700
 Db 7775 TGGAA-----GGACAGGTCTATTACTAAAGCTCCACGCCAGCATCCGAGAGAAAGCAG 7828
 QY 7701 CCAAGAGTTTGGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGGCTTGAAGG 7760
 Db 7829 GTCACTGGTTTGGCCACCAACCGCCCATCACTTGGCAAGGCATCATGTTTGGCATCAAAG 7888
 QY 7761 ATGGCCGAGTACACACAGACATCATCAGTGTGGCCATGAGGATGGCCGAGGTTGCTG 7820
 Db 7889 AAGGGCGGGTGACCAACCGGGGTGTCCAGCATCGCCAGCAAGTAGCCCAAGGTGGCAT 7948
 QY 7821 CCATCTTGAACCATGCGCCACTACCTACAGAGAACCTGCACTTCAACATTGATGGGTGGATA 7880
 Db 7949 CTGTGCTGAACAACCGCTTACTCTGACACAGATGCACTACAGCATCGAGGGCAAGGACA 8008
 QY 7881 CCCATTACTTTGTGAACCAACAGGACCTTCAGAGGTGACCTTGGCCATCCTGGGCTCAGTG 7940

Db 8009 CCCACTACTTTGTGAAGATTGCTCAGCGGATGGCGACCTGCTACACTAGGACACCA 8068
 QY 7941 GGGGGCGCGCAACCCCTGGAGAATGGGGTCAACGTCACTGTGTCCAGATCAACACAGTAC 8000
 Db 8069 TCGGCGCGCAAGGTGTAGAGAGCGGGTGAACGTGACCGTGTCCAGGCCACCGTGTCTG 8128
 QY 8001 TTAATGGCAGGAGTACAGCTACACAGACATCCAGTCCAGTACCGGGCAGTGTCTTGA 8060
 Db 8129 TCAACGGCAGGAGTCAAGGTTTCAAGATTTGATTTCCAGTACTCCACGCTGTCTCTCA 8188
 QY 8061 ACACACGCTTACGG-----GACACGTTTGGATGAGGAGAAAGCAGCGGTCTCTGGAGC 8111
 Db 8189 GCATCGCTATGCTCACCCTCCGACACCTTGGACGAAGAGAGGCCCGCTCTGGAGC 8248
 QY 8112 TGGCCCGGACAGAGCCGTGCGCAAGCGTGGGCGCGGAGCAGCAGAGACATGCGGGAAG 8171
 Db 8249 AGCGCAGACAGAGGCGCTTGGGACGCGCTGGGCAAGGAGCAGCAGAAAGCCAGGAGC 8308
 QY 8172 GGGAGGAAGCGCTCGGGCTGGACAGAGGGGGAGAGCAGCAGGTGTGAGCAGAGGC 8231
 Db 8309 GGAGAGAGGGGAGCGCTGTGGAATGAGGGCGAGAGCAGCAGCTTCTGAGCAGCGGC 8368
 QY 8232 GGGTCAAGGCTACGACGCTTTTCTGTGATCTCTGTGAGCAGTACCCAGAACTGTGAG 8291
 Db 8369 GGGTCAAGGCTACGAGGATATTACGTCTTCCCGTGGAGCAATATCCAGAGCTTGCAG 8428
 QY 8292 ACAGCGCCCAACAACTCCACTTTCATGACAGAGAGAGATGGCGGAGGTGACAGA 8348
 Db 8429 ACAGTAGCAGCAACATCCAGTTTTTAAGACAGATGAGATGGGAAAGAGGTAAACAA 8485

RESULT 9
 ADB32023
 ID ADB32023 standard; cDNA; 9826 BP.
 XX AC ADB32023;
 XX DT 04-DEC-2003 (first entry)
 XX DE Human FCTR3b cDNA.
 XX KW Human; gens; ss; FCTR3b; colorectal cancer; adenomatous polyposis coli;
 KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
 KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
 KW renal cell carcinoma; melanoma; clear cell carcinoma;
 KW granular cell carcinoma; neurological disorder;
 KW neurodegenerative disorder; nerve trauma;
 KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
 KW Gardner syndrome; mental health condition; immunological disorder;
 KW allergy; asthma; lung disease; reproductive disorder; deafness;
 KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
 KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
 KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
 KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
 KW lattice corneal dystrophy.
 XX OS Homo sapiens.
 XX PN US2003087816-A1.
 XX PD 08-MAY-2003.
 XX PP 05-MAR-2001; 2001US-00800198.
 XX PR 03-MAR-2000; 2000US-0186592P.
 XX PA (VERM/) VERNET C.
 PA (FERN/) FERNANDES E.
 PA (SHIM/) SHIMKETS R.
 PA (HEER/) HERRMANN J.
 PA (MAJU/) MAJUMDER K.
 PA (MACD/) MACDOUGALL J.
 PA (MISH/) MISHRA V.

Db 1739 GGAGCATACAGACCTTGGTTGAGATGAAGCCGTTGTTGTGAGTACCTGAGTGGGCC 1798
 QY 1647 TGTGGCACTTGGCTTTTACATGACGGAAGAGTGCAGAGTGGTTTCTCTTCTACCA 1706
 Db 1799 TGTGGCATCTGGCTTCTACATGATGGAAGAGCAAGAGATGGTTCTCTTCAATCTG 1858
 QY 1707 CTGSCATTGAGTCGCTGATTAACCTCCAGCACTGCTATGGAATGGTCACTGCTCT 1766
 Db 1859 TTGTCCTAGATTCACTGAGCACTGTCACGTAATCTGCAATGGGAATGGTGAATGTGT 1918
 QY 1767 CTGGACCTGCCACTGCTCTCTGGGTTTCTGGGCCCGGACTGTGGCAGAGCCTCTGCC 1826
 Db 1919 CCGGGTGTGTCCTGTTTCCAGGATTTCTAGGAGCAGACTGTGCTAAAGCTGCTGCC 1978
 QY 1827 CCGTCTCTGTAGCGGAATGGCCAAATACATGAAGGAGATGCTTGTGCCACAGTGGCT 1886
 Db 1979 CTGTCTCTGTGAGTGGGAATGGCAATATTCTAAGGAGCAGTGGCAGTGTACAGCGCT 2038
 QY 1887 GGAAAGGCGCTGAGTCGATGTGCCCAACCAACCACTGATCCATGTGCTGCGCAAC 1946
 Db 2039 GGAAAGGTCAGAGTGCAGTGGCCATGATCACTGATCATGATCTCTGCGGGGCC 2098
 QY 1947 ATGGCACTGTATCAAGGCACTGCTGATCTGCAACCTTGGCTACAAGGGGAGAGTGTG 2006
 Db 2099 ACGGCTCTGATGATGGGAATCTGTGCTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2158
 QY 2007 AGGAAGTGGACTGTGATGACCCCATGTTTCAGGCGGGGTGCTGGTGTGAGAGCGCAAT 2066
 Db 2159 AGGAAGTTGATGCTGTGATCCACCTGCTCCAGCCAGGAGTCTGTGATGGAGAA 2218
 QY 2067 GCCATTGTTTGGGATGGGAGGACCAACTCCGAGACCCCGAGGCCACATGCTTAG 2126
 Db 2219 GCTGTGAGCGCTGGCTGGGTGCTGTAACCTGTGAGTGGCGAGGCTCAGTGGCCAG 2278
 QY 2127 ACCAGTGTTCAGGCAAGCACTTCTCCCGGACACCGGGCTTTCAGCTGTGACCCAA 2186
 Db 2279 ACCAGTGTGGGATGGGACGACCTGCTGACCGGACCGGCTTTCAGCTGTGATGCTG 2338
 QY 2187 GCTGACTGTGACAGCTGTTCTATCGAGATCTGTGCTGCGACTGTGCTGGCCATGGCG 2246
 Db 2339 ACTGATGGTCCGACTGCTCTGTGAGTTGTGCTCAGTAGACTGTGGCACTCAGCG 2398
 QY 2247 TGTGGTATGGGGCACCTGCGCTGCGAGGATGCTGGATGGGGGACGCTTGGCCACGC 2306
 Db 2399 TCTGATCGGGGAGCTGCGCTGTGAAGAGGCTGGACAGGCGCAGCGTGTGACCGC 2458
 QY 2307 GGGCTGCGACCGCGCTGTCGAGATGGGACCTGCGCGAGCGCAAGTGGCAGTGCA 2366
 Db 2459 CGGTGTGCAACCCCGCTGATGAGCAGCGGACCTGTAAAGATGGCAATGTGAATGCC 2518
 QY 2367 GCGTGTGGAATGGCGAACTGTGACCACTGCTACTATCTGATAGGATAGGATTAAG 2426
 Db 2519 GAGAGGGCTGGAATGGTGAACACTGCACTGCTGTTGGTAGGCAAAACGCGAGGACCGAAACAG 2578
 QY 2427 AGGTTGCGCTGGTGTGCAATGGCAACGCACTGCTAGACTGATGCTGATGCTGGTGGC 2486
 Db 2579 ATGCTGCGCTTGAATGTGCAACGGTAAAGGGATGCACTGGGTGAGAAACAGCTGGC 2638
 QY 2487 ACTGCTGCTGCGCTGAGAGAGTGGCTGTGACACTTCCATGAGAGCTGCCT 2546
 Db 2639 AGTGTGTCTGCCAGCGCTGGAGAGGCGCGGATGCAACGTTGCCATGGAATTCCT 2698
 QY 2547 GCGGTGACAGCAAGACATGATGGAGTGGCTGTGGTGTGATGATGAGACCTGCTGCT 2606
 Db 2699 GTGCTGTAAACAGGATTAAGGGAGATGGCTGTGGATGTTTGGACCTGACTGCT 2758
 QY 2607 GCTCCAGCGCTGTGCAATATCAACCGGTGTGCTTGGCTGCTCCCTTAAACCTGTGACA 2666
 Db 2759 GCCTGCACTGAGCTGTGACAGACGCTGCTGCTGCGGGGGTCCCGGACCCACTGGACA 2818
 QY 2667 TCATCCAGGACACAGGTCTCTGTGTCACAGCGAAACCTTACACTCTCTTCTATGACCGCA 2726

Db 2819 TCATTTCAG-----CAGGGCCAGACGGATTGGCCGCGAGTGAAGTCTTCTATGACCGTA 2872
 QY 2727 TCAAGTTCCTGTGGCAGGACAGACGCAATTAATCCCGGGAGAACCCCTTTGATG 2786
 Db 2873 TCAAGCTCTGGCAGCAGGATAGCACCAATCATCTCTGGAGAGAACCTTTTCAACA 2932
 QY 2787 GAGGGCATCTTGTGTATTCTGTGGCCAGTGCATGACATCAGATGGAACCCCTTGTG 2846
 Db 2933 GAGCTTGTGTTCTCTATCCGAGCCAGTGTAACTACAGATGGAATCCTCCCTGGTGG 2992
 QY 2847 GTGTGAACATCAGTTTTGTCAATAACCCCTCTCTTTGGATATACATCAGCAGGCAAGATG 2906
 Db 2993 GTGTGAACGTCTTTTGTCAAGTACCCAAAATACGGCTACACCATCACCGCCAGGATG 3052
 QY 2907 GCAGCTTTGACTTGTGACAAATGCGGATCTCCATCATCTGCGGTTTCAGCGGGCAC 2966
 Db 3053 GCAGCTTGCATGATGCAATGAGGTGCTTCTTCACTACACTTTGAGCGAGGCC 3112
 QY 2967 CTTTCATCACAGAGACACCCCTGTGCTGCCATGGATCGCTTCTTTGTCATGGA 3026
 Db 3113 CGTTTCATGAGCCAGGAGCGCACTGTGTGGCTGCCGTGGAAACAGCTTTACGCCATGACA 3172
 QY 3027 CCATCATCATGAGACATGAGGAGATGATTCCTCAGCTGTGACCTGAGCAATTTTGGCC 3086
 Db 3173 CCGTGTGATGAAGACCCAGGAGAACTCCATCCCGAGCTGTGACCTCAGTGGCTTTGTCC 3232
 QY 3087 GCGCAACCCAGTCTCTCCATCCCGACAGCTCTTCCGACAGCTCTGTGCGAGAGA 3146
 Db 3233 GGCCTGATCCATCATCTCTCCACAGTGTCCACTTCTTTAGTGTGCTGCCCCGTGGC 3292
 QY 3147 AAGGCCCATTTGTCCGGAATTCAGGCTTTTCAGAGGAAATCTTATCTCTGCTGCA 3206
 Db 3293 AGAATCCCATCTGCTGACAGCCAGGTTCTTCATGAAGAAATCAGCTCCCTGGTCCA 3352
 QY 3207 AGATGAGCTGAGCTACCTGAGCAGCGGACCCCTGGCTGCAATCTGCTCTGAGATCA 3266
 Db 3353 ATGTGAATCTGCTGATCTGAGCTCTAGAATGCGAGGTCAAGTCACTGTCTGAAGATCA 3412
 QY 3267 GCGTCAACCCAGCCATCCCTTCAACCTCATGAAGTGCACCTCATGTGTAGCGGTGG 3326
 Db 3413 CCATGACCCAGTCCACAGTGCCTTCAACCTCATGTAGGTTCCACCTGATGTGTGCTG 3472
 QY 3327 AGGGCGGCTCTTTCAGGAAGTGGTTCGCTGAGCGCCAGACCTGTCTTATTTTCAATT 3386
 Db 3473 AGGGCATCTCTTCCAGAGTCAATCCAGGCTTCTCCCAACCTGCGCTCCACCTTCACT 3532
 QY 3387 GGGACAGACAGAGCTCTACAAACAGAGGTTTGGGCTTTCAGAACCTTGTTCCTG 3446
 Db 3533 GGGACAGACAGATGCTATGGCCAAAGGTTGTATGACTCTCAGATGCTGTGTGCTG 3592
 QY 3447 TGGTTATGAATATGAATCTTCCCGAGATCTAATCTCTGTGGAAAAAAGAACAGTGC 3506
 Db 3593 TCGGTTTGAATATGAGACCTTCTCCAGTCTAATTTCTCTGGGAGAAAGGACAGCCCTCC 3652
 QY 3507 TCGAGGCTATGAATTCAGCGCTCCAGCTTGGAGGATGAGCCTAGACAAACATCATG 3566
 Db 3653 TTCAGGATTCAGCTGGACCCCTCCAACTCGGTGGTGTCTCTAGACAAACACCA 3712
 QY 3567 CCTCAACATTCAGAGTGTATCTTCGCAAAAGGAAATGGGGAAACAGTTTGTGTCTC 3626
 Db 3713 TCCTCAATGTTTAAAGTGGATCTTACAAAGGCACTGGGGAAAAACAGTTCTCTGACCC 3772
 QY 3627 AGCAGCTCTGTGCTATGGGAGCATCATGGCAATGGCGCGGAGAGCATCTCTGCTGCC 3686
 Db 3773 AGCAGCTTGCATCATCAGGCAATGCTGGCAATGCTGCGCGGAGGCAATTTCTCTGTC 3832
 QY 3687 CCAGCTGCAAGCGGCTTCTGACCGCAACAAAGCTCTTGGCCCCCAGTGGCCCTCACTGTG 3746
 Db 3833 CCAGCTGCAACGGCTTCTGAAGGCAACAAAGTGTGCGCCCCAGTGGCTCTGGCTGTG 3892
 QY 3747 GCTCTGACGGGAGCTCTATGTGGTGTATTTCAACTACATGAGAGATCTTCCCTCTG 3806
 Db 3893 GAATCGATGGGAGCTCTATGTGGTGTGACTTCAATTTACATCCGACGATCTTCTCTCTC 3952

QY 3807 GAATGTACCAACATCTAGAGCTGAGGATTAAGATTTCAGACATAGTCAAGTCCAG 3866
 Db 3953 GAATGTACCAACATCTAGAGCTGAGGATTAAGATTTCAGACATAGTCAAGTCCAG 4012
 QY 3857 CACACAAATACATCTAGAGCTGAGGATTAAGATTTCAGACATAGTCAAGTCCAG 3926
 Db 4013 CACACAAATACATCTAGAGCTGAGGATTAAGATTTCAGACATAGTCAAGTCCAG 4072
 QY 3927 ACAGCGCGGGTCTTTAAATCAAGTCCATCTGAGTGGTGAAGGACCTTGTCAAGT 3986
 Db 4073 ACAGCGCGGGTCTTTAAATCAAGTCCATCTGAGTGGTGAAGGACCTTGTCAAGT 4132
 QY 3987 CTGAGGTGCTGAGGACAGGTCACAGTCCCTCCCTTGTGATGACATCTGCTGCGGG 4046
 Db 4133 CGGAGTGTGCGGACAGGTCACAGTCCCTCCCTTGTGATGACATCTGCTGCGGG 4192
 QY 4047 ATGTTGGGAAGCCACAGAGCCACATCAACATCCAGGGGTATTACAGTGGACAGT 4106
 Db 4193 ATGAGGGAAGCCATAGATGCAACCTGATGAGCGCGGAGAGGTATTGCGAGTACACA 4252
 QY 4107 TTGGGCTGATCTAGTTCGTGATGCGACATGATCAGACGATCAGATCAGATCGGATCA 4166
 Db 4253 ATGGGCTCATGTTCTGATGCGACATGATCAGACGATCAGATCAGATCGGATCA 4312
 QY 4167 TCTCACCTGCTCGGCTCTAATGATCTACATCAGCCCGGCCACTCAGCTGATCTG 4226
 Db 4313 TCTCACCTGCTCGGCTCTAATGATCTACATCAGCCCGGCCACTCAGCTGATCTG 4372
 QY 4227 TCATGGATATTCCAGTAGAGTGGAGTGGCCACAGATCAGATCAGATCAGATCGGAT 4286
 Db 4373 GCATGGATATTCCAGTAGAGTGGAGTGGCCACAGATCAGATCAGATCAGATCGGAT 4432
 QY 4287 ACACTCATCTTATGCTCTGACACAAATGTTGCTGCAATCTGCAAAACACACAGG 4346
 Db 4433 ATACTCTTGTATGTTCTAGAGACATGTCATCTGCAATCTGCAAAACACACAGG 4492
 QY 4347 TGGCATTTGCTGCGGAGGCGCCATGATCTGCGAGTCCCTGCGATTTGATCAGATCT 4406
 Db 4493 TCAGCATATTGCGGAGGCGCCATGATCTGCGAGTCCCTGCGATTTGATCAGATCT 4549
 QY 4407 TAAGCAAGTGGCCATCAGCAACCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 4466
 Db 4550 TCAGCAAACTAGCATCTGCTGCGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 4609
 QY 4467 ATGGGCTCTGATATGCTGAGACTGATGAGAAAGATCAACCGCATCAGGAGTCA 4526
 Db 4610 CTGGGCTCTTACATCTAGACAGATGAGAGATTAACCGCTACGCGCAGGTAA 4669
 QY 4527 CCACTAGTGGAGATCTCAGTCTGCTGGGCGCCAGTGGCTGATGATGATGATGATG 4586
 Db 4670 CAACCAACGGGAGATCTGCTTTAGTGGGCGAGCTCGGATGCGAGTGGAGTGGAGT 4729
 QY 4587 ATGCCAAGTGTATGTTTCTGGAGACGATGTTATGCCAAGGATGCAAGTAAATA 4646
 Db 4730 ATGTCAATGCAATGCTATTACAGAGATGATGCTACGCGATGATGCTGATGAT 4789
 QY 4647 CCCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4706
 Db 4730 CCCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4849
 QY 4707 TCCGATTCGTTTATCCGAGAGACAGCTTTCTCAACACCCAGACATGATGAGC 4766
 Db 4850 TTCGATCAGGCGGTACAGAGACAGCTTTCTCAACACCCAGACATGATGAGC 4909
 QY 4767 TGTCTTCAACATTCACAGAGTCTTATCTGTTGATGATGATGATGATGATGATGATG 4826
 Db 4910 CTGCATCCCGGAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATG 4969
 QY 4827 CCAAGCTGCCACAGAGACTCTGATCACTTCACTTCACTTCACTTCACTTCACTTCA 4886
 Db 4970 CTGTAGGCTGTGACAGGAGTACTTGTACAAATTTACATATAGTACTGACAAATG 5029

QY 4887 TCACATCTCATCAGACAAATGGCAACATGGTAAATGTCCGCGAGACTCTACTGGGA 4946
 Db 5030 TCACATCTCATCAGACAAATGGCAACATGGTAAATGTCCGCGAGACTCTACTGGGA 5089
 QY 4947 TGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5006
 Db 5090 TGGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5149
 QY 5007 GTGCACTCAAGAGTGTGACACACAGGACAGGTTGGCCATGATGACATACCATGGCA 5066
 Db 5150 GAGGCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5209
 QY 5067 ATTCGGGCTCTTGGCAACCAAAAGCAATGAAACCGATGGAACAAATTTATGATGAG 5126
 Db 5210 ACAGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5269
 QY 5127 ACAGTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5186
 Db 5270 ACCAGGAGGCGCTGACCAAGTGGCGGCCCGGCGGCTGGAACAGTCTGTCACC 5329
 QY 5187 GTGATACAGACAGTTCAGTGCATGTCAGGTAGAGACCTTCCA--GCAAGGATGATGCA 5243
 Db 5330 GGAATGAGAAATCTATTACATTTGACATTTGAGAACTCCAAACCGTATGATGATGCA 5389
 QY 5244 CCATAACCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5303
 Db 5390 CTGTCTATCAACACCTCTCTTCTGATGAGGCTCTCTACACAGTGTACAGATCAAGTTC 5449
 QY 5304 GGAACAGCTACTACATCGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5363
 Db 5450 GGAACAGCTACTACATCGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5509
 QY 5364 AGTGGCGCTGACAGCTGAGCCCCACTTCTGCTGGTGGCAACCGTCAACCCACCGTGGGCA 5423
 Db 5510 GTATCAGCTTCCACAGGAGCCCCATGCTTAGCGGCGCACCATCACCCACCATTTGGAC 5569
 QY 5424 AGAGGATGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5483
 Db 5570 GCTGCAACATCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5629
 QY 5484 AAGAGAGGCTCGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5543
 Db 5630 AGGAAAGATTAAGGCAAGTCAACCATCTTTGGCAGGAGCTCCGGTCCATGGAAGAA 5689
 QY 5544 ATCTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5603
 Db 5690 ATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5749
 QY 5604 GCAAGTTCACCTTCGGATCTGTACGACAGCGGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5663
 Db 5750 GGAAGTTCACCTGAGGATCAATTTATGACAGGTTGGCGCGGCTTCTCTGCTGCTGCTGCTGCTGCTG 5809
 QY 5664 GCAGAGGCTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5723
 Db 5810 GCAGCGGCTGGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5869
 QY 5724 AGAGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5783
 Db 5870 AGCGTGGGCTGAGCGGAGGACAGATCATGCAAGCAAGCGCCGCTGCTGCTGCTGCTGCTGCTGCTG 5929
 QY 5784 TCTTCTGCTGATGGAAGACATGGAGCTACACATCTTAGAGAGTCCATGGTGTCTCTAC 5843
 Db 5930 TGTTCGCTGACGGAAGTGTGGAGTCTCTTACCTTGAAGTCCATGGTCTCTCTGCTGCTGCTGCTGCTGCTG 5989
 QY 5844 TACACAGCGAGGAGTATCTTTGAGTTCGACAGAGTATGACCGCTCTCTTCTGCTG 5903
 Db 5990 TCCAGAGCAACGCTGATGATATTTGAGTATGACTCTCTGACCGGCTCTCTTGGCGTCA 6049
 QY 5904 CGATGCCCAACGCTGCGGCGGACACTAGAGACCTCCGCTCAGTGGGCTACTACAGAA 5963
 Db 6050 CCATGCCCGGCTGCGGCGGACAGATGTCCACACACACCTCCATCGGCTACATCGTA 6109
 QY 5964 ACATCTATAGCCCCCTGAGGGCAATGCTCTAGTCTATACAGGACTTCACTAGGATGGGC 6023

Db 6110 ATATTACAAACCGCCTGAAGCAATGCTTCGGTCACTTTGACTACAGTGATGACGGCC 6169
QY 6024 ACCTCTTTCACACCTTCTACCTGGGCATCTGGCGGAGGTGATATACAGTATGGCAAC 6083
Db 6170 GCATCCTGAAGACCTTCTTTTGGGCACCGGACCGCAGGTGTTCTACAAGTATGGAAAC 6229
QY 6084 TGTCAAAGCTGGCAGAGCGCTCTATGACACCAACCAAGGTCACTTCACTTATCAACGAGA 6143
Db 6230 TCTCAAAGTTATCAGAGATTCTACGACAGTACCGCGCTCACCTTCGGGTATGACGAGA 6289
QY 6144 CGGAGGCAATGCTGAAGACCAATCAACCTACAGATGAGGGGTTCACTGACCAATCCGCT 6203
Db 6290 CCACTGGTGTCTTGAAGATGCTCAACCTCCAAAGTGGGGGTTCTCCTGCAACATCAGGT 6349
QY 6204 ACCGTACAGTGGGCGCCTGATTGACCGACAGATCTTCCGTTCACTGAGGAGGCAATGG 6263
Db 6350 ACCGGAAGATTGGCCCGCTGTGGACAAAGCAGATCTACAGTTCTCCGAGGAAGCATGG 6409
QY 6264 TCAACGCCCGTTTGACTACAACCT---ATGACAAACAGCTTCCGGGTGACCAGCATGCAGG 6320
Db 6410 TCAATGCCAGGTTTGACTACACCTATCATGACAAACAGCTTCCGCATCGCAAGCATCAAGC 6469
QY 6321 CTGTGATCAACGAGACCCCATGCGCCATTGATCTCTATCGTATGATGATGTGTGAGCA 6380
Db 6470 CCGTCAATAGTGAGACTCCCTCCCGGTGACCTTACCGCTATGATGAGATTTCTGGCA 6529
QY 6381 AGACAGACAGTTTGGGAAGTTTGGTGTCACTTACTATGATTAACCAACAGATCAACCA 6440
Db 6530 AGGTGGAACACTTTGGTAAGTTTGGAGTCACTATTATGACATCAACAGATCATCACCA 6589
QY 6441 CAGTGTGATGACCAACCAAGCAATTTGATGCAATATGGCAGATGAAGAGTGCAGT 6500
Db 6590 CTGGCGTATGACCTTCAGCAAAACACTTCGACACCCATGGCGGATCAAGGAGTCCAGT 6649
QY 6501 ATGAGATCTTCGGTTCGCTCATGTACTGATGACCGCTCCAGTATGATAAATGGGGGAG 6560
Db 6650 ATGAGATGTTCCGGTCCCTCATGTACTGATGACGGTGCATATGACAGCATGGGCAGG 6709
QY 6561 TAGTGAAGAGAGCTGAAGGTAGGACCTTACGCCAATACCACTTCGCTACTCTCTATGAT 6620
Db 6710 TGATCAAGAGGAGCTAAACCTGGGGCCCTATGCCAATACCAAGATACACCTATGACT 6769
QY 6621 ATGATGCTGACGGCCAGCTGACAGACTCTCCATCAATGACAGCCACTCTGGCGCTACA 6680
Db 6770 ACGATGGGAGCGGAGCTCCAGAGCGTGGCCGTCAATGACCGCCGACCTGGCGCTACA 6829
QY 6681 GCTACGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGGGAAACAGTGCACGGCTCACAC 6740
Db 6830 GCTATGACCTTAATGGGAATCTCCACCTTACTGAAACCCAGGCAACAGTGTGGCTCATGC 6889
QY 6741 CACTACGATGATACATCGGACCGCATCTCGCTGGGTGAGTGCAGTGCATACAGATGG 6800
Db 6890 CTTTCGCTATGACCTCGGATCGGATACCAAGCTCGGGGATGTGCAGTACAAAATTG 6949
QY 6801 ATGAGATGGCTTCTGAGGAGCGGGCGGTGATATCTTTGATACAACTCAGCTGGCC 6860
Db 6950 ACGACGATGCTATCTGTGCAGAGAGGGTCTGACATCTTCGATATCAATTCGAAGGCC 7009
QY 6861 TGCTCATCAAGGCTTACAAACGGGCTGGCAGTGGAGTGCAGTACCGCTACGATGGCC 6920
Db 7010 TCCTAACAAAGAGCCTCAACAAAGCCAGCGGTTGGAGTCCAGTACCGCTATGATGGCG 7069
QY 6921 TGGGCGGCGGTGTGCAGCAAGAGAGCCACAGCCACCACTTCAGTCTCTCTATGCGAG 6980
Db 7070 TAGACGGCGGCTTCTCAAGACCAACCTTGGGCCACCACTTCAGTACTTCTCTCTG 7129
QY 6981 ACCTGACCAACCCCAACCAAGTCAACCACTGTACAAACCACTCCAGCTCTGAGATCACCT 7040
Db 7130 ACCTTCCCAACCCGACCGCATACCCATGTCTACAACTCACTCCAACTCGGAGATTACCT 7189
QY 7041 CCCTTACTACGACTTGCAGGACACCTCTTGCATGGAGCTGAGCAGTGGTATGAGT 7100

Db 7190 CACTGTACTACGACCTCCAGGGCCACCTCTTTGCCATGGAGAGCAGTGGGGAGGAGT 7249
QY 7101 TTATCATAGCTTGTGACAAACATCGGACCCCTCTTCTGTCTTTAGTGGAAACAGGTTTGA 7160
Db 7250 ACTATGTTGCCCTCTGATACACAGGAGCTCTCTGGCTGTGTTCAGCATCAACGGCTCA 7309
QY 7161 TGATCAAGCAAAATCCTGTACACAGCTATGGGAGATCTATCATGATACCAACCCCAACT 7220
Db 7310 TGATCAACAGCTGACAGTACACGGCTATGGGAGATTATATGACTCCAAACCCGACT 7369
QY 7221 TTCAGATCATCATAGCTACCATGTGTGGCTCTATGATCCACTCAACAGCTTGTTCACA 7280
Db 7370 TCAGATGTCTATGGCTTCCATGGGGACTCTATGACCCCTGACCAAGCTGTGTCTCACT 7429
QY 7281 TGGGCGGCGGAGATTATGATGTGTGGCGGAGCGTGGACTAGCCGACGACCAAGCTGT 7340
Db 7430 TCACTCAGCGTGTATGATGATGCTGGCAGGAGTGGACCTCCGACACTATACCATGT 7489
QY 7341 GGAAGCACCTTAGTAGCAGCAACGTCTATCCCTTTTAAATCTCTATATGTTCAAAAACAACA 7400
Db 7490 GGAATAACGT---GGGCAAGGAGCGCGCCCTTTAACTGTATATGTTCAAGAGCAACA 7546
QY 7401 ACCCATCAGCAACTCCGAGGACATCAAGTGTCTTATGACAGATGTTTAAACAGCTGGCTG 7460
Db 7547 ATCCTCTCAGCAGTGAAGTATGTAAGAACTACGTGACAGATGTGAAAGAGCTGGCTTG 7606
QY 7461 TCACCTTTGGATTCCAGCTACACAACTGATCCCTGGTATTCCCAAAACAGACATGATG 7520
Db 7607 TGATGTTTGGATTTCAGCTTAGCAACATCAITCTCTGGCTTCCGAGAGCCAAATGATTT 7666
QY 7521 CCATGGAACCTTCTACAGCTCATCCACACAGATGAAACCGCAGAGTGGGACACA 7580
Db 7667 TCGTGCCTCTCCCTATGA-----ATTGTACAGAGTCAAGCAAGTGAATG 7714
QY 7581 GCAAGTCTATCTCGGGGTCAAGTGAAGTACAGAGCAGCTCAAGGCTTTGTCACT 7640
Db 7715 GACAGCTATTACAGTGTCCACAGACAACAGAGACATAACAGGCTTTCATGGCTC 7774
QY 7641 TAGAAGCTTTGACCGACTCTATGGCTCCAATCACAGTGCACAGCTCCAGGCTCCAAAGA 7700
Db 7775 TGGAA-----GGACAGGTCTATTAAAGAGCTCCAGCCAGCATCCAGAGAAAGCAG 7828
QY 7701 CCAAGAGTTTGCATCCAGCGCTCAGTCTTTGGCAAGGGGTCAAGTTTCCCTTGAAG 7760
Db 7829 GTCATGTTTGGCCACCAACAGCCCTCATTTGGCAAGGATCATGTTTCCCTCAAG 7888
QY 7761 ATGGCCGAGTACCAACAGACATCATGTTGGCCATGAGGATGGGCAAGGTTGCTG 7820
Db 7889 AAGGCGGCTGACCAACGCGGTGTCCAGCATCGCCAGGAAAGTAGCCGCAAGTGGCAT 7948
QY 7821 CCATCTTGAACCATGCCACTACCTAGAGAACCTGCATTCACCTTACCATGATGGGTGAT 7880
Db 7949 CTGTCTGAACAACGCTTACTCTGGAACAAGTGCATCAGCATCGAGGCAAGGACA 8008
QY 7881 CCCATTTACTTTGTGAACCAAGACCTTCAGAAAGTGAACCTGCGCATCTCTGCGCTCAGT 7940
Db 8009 CCCACTACTTTGTGAAGTTTGGCTCAGCCGATGGCGACTTGTCACTACGACCAACA 8068
QY 7941 GGGGCGGCGGAACCTCGAGAAATGGGTCACAGTCACTGTGTCCAGATCAACACAGTAC 8000
Db 8069 TCGGCGCAAGGTGTAGAGAGCGGGTGAACGTGACCGTGTCCAGCCCACTGCTGG 8128
QY 8001 TTAATGGCAGGACTAGACGCTACACAGATCCAGTCCAGTCCAGTGGGCACTGTGCTTGA 8060
Db 8129 TCACGCGAGACTCGAAGTTTCAGAACATTTAGTTTCCAGTACTCCACGCTGTGCTCA 8188
QY 8061 ACACACGCTACGG-----GACAAAGTTGATGAGGAAGGCAAGGCACTGCTGAGC 8111
Db 8189 GCATCCGCTATGGCTCACCCCGACACCTTGGACGAAGAGAGGCGCGCTCTCTGGACC 8248
QY 8112 TGGGCGGCGCAGAGCGCTGCGCAAGCTGGGCGCGGAGCAGCAGAGACTGCGGGAAG 8171
Db 8249 AGGCGAGACAGAGGCGCTGGGCAAGGCTGGGCAAGGACAGCAAGAACCCAGGACG 8308

QY 8172 GGGAGGAGGCTCGGGCTGGACAGAGGGGGAGAGCAGCGTCTGAGCACAGGC 8231
 DB 8309 GGAGAGGGAGGGCTGGACTGGAGTGGAGGAGAGAGCAGCTTCTGAGCAGCGGC 8368
 QY 8232 GGGTGCAGAGGTACGAGCGCTTTTTCGTGATCTCTGTCGACGAGTACCCAGAACTGTCAG 8291
 DB 8369 GCGTGAAGGTACGAGGATATTACGTGCTTCCGTGGAGCAATACCCAGAGCTTCAG 8428
 QY 8292 ACAGGCCCAACATCCACTTCATGACAGAGCGAGTGGCGGAGGTGACAGA 8348
 DB 8429 ACAGTAGCAGCAATCCAGCTTTTAAACACAGATGATGGGAAAGAGTAACAAA 8485

RESULT 10
 AAS14089
 ID AAS14089 standard; DNA; 9729 BP.
 AC AAS14089;
 XX
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human FCTR3f DNA sequence.
 XX
 KW Human; FCTR3; myelogenous leukaemia; carcinoma; melanoma; glioma; ds;
 KW astrocytoma; congenital neonatal alloimmune thrombocytopenia; infection;
 KW neurological disorder; neurodegenerative disorders; nerve trauma;
 KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
 KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
 KW mental health condition; immunological disorder; allergy; infertility;
 KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
 KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
 KW demoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
 KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
 KW Spino cerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
 KW Corneal dystrophy-Greonow type I; Corneal dystrophy-lattice type I;
 KW Reis-Bucklers corneal dystrophy; cystostatic; immunosuppressive;
 KW anti-allergic; antiasthmatic; antinfertility; antiinflammatory;
 KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
 KW gynaecological; antinfertility; immunostimulant; auditory; haemostatic;
 KW gene therapy; FCTR3a; neurestin-like protein; FCTR3f.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT 5'UTR 1..209 /tag= b
 FT CDS 210..8384 /tag= a
 FT 3'UTR /product= "Human FCTR3f"
 FT 8385..9729 /tag= c
 XX
 XX
 PN WO200156747-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001WO-US007160.
 XX
 PR 03-MAR-2000; 2000US-0186592P.
 PR 03-MAR-2000; 2000US-0186718P.
 PR 06-MAR-2000; 2000US-0187293P.
 PR 06-MAR-2000; 2000US-0187294P.
 PR 17-MAR-2000; 2000US-0190400P.
 PR 07-APR-2000; 2000US-0196018P.
 PR 03-JAN-2001; 2001US-0259548P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
 PI Maccougall J, Mishra V, Mezes PS, Rastelli L;
 XX
 DR WPI; 2001-596837/67.

DR P-PSDB; AAU08681.
 XX Novel polypeptides designated as FCTR3 polypeptides, useful in detection,
 PT prevention and treatment of a broad range of pathological states.
 XX
 XX Claim 9; Page 37-39; 215pp; English.
 XX
 CC The invention relates to human FCTR3 polypeptides, FCTR1-FCTR7, and the
 CC nucleic acids encoding them. These sequences are useful for the treatment
 CC or prevention of numerous disorders including myelogenous leukaemia,
 CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
 CC alloimmune thrombocytopenia, neurological disorders, neurodegenerative
 CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
 CC -Tooth neuropathy, demyelinating Gardner syndrome, familial
 CC myelodysplastic syndrome, mental health conditions, immunological
 CC disorders, allergy and infection, bronchial asthma, Avellino type
 CC eosinophilia, lung diseases, reproductive disorders, infertility, male
 CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
 CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
 CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
 CC infection Spino cerebellar ataxia, Plasmodium falciparum parasitaemia,
 CC Corneal dystrophy-Greonow type I, Corneal dystrophy-lattice type I and
 CC Reis-Bucklers corneal dystrophy. This sequence represents DNA encoding
 CC the FCTR3a homologue FCTR3f
 XX
 SQ Sequence 9729 BP; 2548 A; 2629 C; 2438 G; 2114 T; 0 U; 0 Other;
 Query Match 40.1%; Score 3351.4; DB 5; Length 9729;
 Best Local Similarity 64.6%; Pred. No. 0;
 Matches 5385; Conservative 0; Mismatches 2771; Indels 181; Gaps 19;
 QY 35 ATGACGCTGAAGAGAGAGAGCTTACCGCTCGTACCGCGCGCGCGCGAGCGC 94
 DB 210 ATGGATGTAAGAGACCGCGGACACCGCTCTTTGACGAGAGGACGCTGTGGCAAGAGTGT 269
 QY 95 CGCTACACAGCTCGTCGCGGACAGCGAGGAGGCAAG---CCCGCAGAAATCGTAC 151
 DB 270 CGCTACCAAGTCTCTCTTGACAGTGTGAGTGTGCGGTGCCACACAGAAATCTTAC 329
 QY 152 AGCTCCAGCGAGACCGCTGAAGCGCTTACGACGAGCGCGCGCTAGCTATGCGAGCGC 211
 DB 330 AGCTCCAGTGAGACTCTGAAGCGCTATGACCATGACGAGGATGCACTATGAAACCGA 389
 QY 212 GTCAAGGACATGTGCGCGAGAGCGCGGAGGAGTCTGCGGCACAGGTCCCACTTACC 271
 DB 390 GTCAAGACCTCATCCACCGGAGTCAATGATGATTTCTTAGAACAAGAACCACTTACC 449
 QY 272 CTGCGGAGCTGGGGCTGGAAGAAGTAACGCCCTTACGCGACCGCTGTACCGGACAGAC 331
 DB 450 CTGCGGAACTGGGCATCTGTGA---GCCCTCCACACCGAGCGGTACTGCTCCGAC 506
 QY 332 ATTGGCTGCCCAATGCGGCTACTCCATGCGGGGTGGCTCTGATGCCGACATGAGGCT 391
 DB 507 ATGGGATCTTTCACGAGGGCTACTCCCTTAGCAGAGGCTGTGACGCGCACTCCGACAC 566
 QY 392 GACACGCTGTGCCCTGAGCACCGCGTCTGTGGGCGCGGAGGACACCGGTCCAGG 451
 DB 567 GAGGAGGATGTCTCCAGAACCGCCATCAAGTGTGGGGCAGAGGATTAATCCAG 626
 QY 452 CGCAGCTCTGCTGTCCAGCGCGGCAATTCATCTCACACTCACCGACCGAGCAT 511
 DB 627 CGCAGTTCGGGCTGTGCGAGTCTGTGAAACTCGGCCCTTACCTGACTGACTGTGACAAC 686
 QY 512 GAAACACTGAGACTGATCATTCGGGGCGGC---TGCAGAACCGCGCGGCTCCGGAC 567
 DB 687 GAAACAAATCATGATGATGAGAACCGGTGCTCCCATTTCCACCTTACCTCTGGCTAGTCTC 746
 QY 568 GCCGCGCGCGCGCTCTCGCAGCGCCACACCGCCCAACAGCAGCAGCGCGCTCCATTAA 627
 DB 747 CTCCTCATCTGCTCAGTGTGCTTCCCATATCTCTCCACAGTGTAGCTGCC-AGATGCC 805
 QY 628 CTCCTGTAACCGGGGCAACTTTCAGCGCGGAGGAGAACCCCGCGCGCGCGCGCGACCA 687

Db 806 ATTGCTAGACAGCAACACCTCCATCAATCATGACACCAACCCCTGATGAGGAATTC 865
Qy 688 CTCGCTCTCCGAGAGCCCTCGCGGGCGCCAGGAGCTGCCACGCCAGAGAA 747
Db 866 CCCCAATTCACTACCTGCTCAGAGATGCTCAGGGCCCCAGCAAGCTCCAGAGTGGCC 925
Qy 748 CTGGTGTCTCAACAGCAACATCCCTCGAGACAGGAACTTAGGCAAGCAGCATTCCT 807
Db 926 TCCGAACACACAGCCAGTGCAGCTCTGAGGCC-----CCTCTCCC 967
Qy 808 AGGACATTGAGAGCAACCTCATTTGATGGAATTTCTGGGCGCTCCCGCCATGATG 867
Db 968 ACCCCCTCACAAACACAGCTGTCCCATCACTCTGTCGCGCAACTCCCTCAACAGAA 1027
Qy 868 GGCCTACAGTGAAGGGCACTTCCTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCAC 927
Db 1028 CTCAGTACCAATCGGCGGAGTCAGATCCACGCCCGCCCGCCAGCCCAATGACCTGGC 1087
Qy 928 CACATCACAGGATCCCACTGACGTCCAGACAGATGATCTCTCCGCGCCGACCCCT 987
Db 1088 CACCACACAGAGTCCGTTCACTTCAGCTTCAGGACAGCTGGGTGCTAAACAGCAACGTGCCACT 1147
Qy 988 GCGCCGAGCACCTTCGCGCGCGGCTTTAACTCTAGAGGCCCTCCAGTACTGTA 1047
Db 1148 GGAGACCGGCACTTCCTTTCAAGACCTCTCTGGGAGCACACCT-----T 1195
Qy 1048 CTGGAAGTGCAGCGCTGAGCGCATGTCATCTCAGCCACTCTGTCATCTCTGCTGGC 1107
Db 1196 GTTCAGCAGCTCTCCCGCGGATACCTTTGACCTCAGGAAGGTTTACAGCGCCCGCC 1255
Qy 1108 ATACTTTGCGCATGCACTGTTTGGCTPAACTGGCACTGCAGCGGATGAGGGGCA 1167
Db 1256 CCGCTGCTGCCAGGAATCTTTCTCCAG-----GAAGCTTTCAA 1297
Qy 1168 GATGATGAGATCAGGAGGACACAGCAGTGGCTGTGCAACCGACGCTTCC-C 1226
Db 1298 GCTGAAGAGCCCTCCAAATCTCAGTGGAAATGTGTCCTCTCCGCAATGCGCG 1357
Qy 1227 TATACCCCTCAGGGGCACTGGCTTAGAGACCCCTGACAGAAAGCAAGCAACACAG 1286
Db 1358 GCGCTCTCTCTGCTATTTGCTGGGTATTTTCATAGTCCCTGCTGTTGAAAAACAG 1417
Qy 1287 AAGGAAAGCCAGTAGTTCTTTCCAGAGGACAGTTTCATAGATTTGGAAGAAATGATG 1346
Db 1418 CAG-----CATAGAGTGGTGAAGCAAG 1443
Qy 1347 TGGGAAGCGAGCTCCAGAGATTCCTCTGSCACTTTCTGAGATCTCAAGTGTCA 1406
Db 1444 TTGTCGCGGGTAAACAAAGAGTCCACCCAGGGGTGTTTGGAGTCAAAATTCACA 1503
Qy 1407 TAGACCATCTGTCATCTGAATTCATGTCTCTGGGAAGGAGCCCTGGTTGGCA 1466
Db 1504 TCAGTCAGCCCACTCTTAAAGTTCAACATCTCCCTCGGGAAGGAGCTCTCTTTGGTG 1563
Qy 1467 TTTATGGCAGAAAAGGCTCCCTCTTCAATACACAGTTTGAATTTGAGCTGCTGG 1526
Db 1564 TTTACATAAGAAGAGGACTTCCACCATCTCATGCCAGTATGACTTCATGAAAGCTGCG 1623
Qy 1527 ATGCGAGAGCTCTTAACCCAGAGGCGGAGCTTAGAGGGAACCCCGCCAGTCTC 1586
Db 1624 ACGGGA-----AGGAGAAGTGGAGTGTGGTTGAGTCTCCAGGGAACGCC 1668
Qy 1587 GGGGAATGTGCCCCCTCCAGCATGAGACAGGCTTCATCCAGTATTTGGAATCAGGAA 1646
Db 1669 GGACATACAGACTTGGTTGAGATGAGCCGTGTTGTGAGTACCTGGATGTGGCC 1728
Qy 1647 TCTGGCACTTGGCTTTTACAATGACGGAAGAGTCAAGAGTGGTTTCTTTCTCAACA 1706
Db 1729 TGTGGCATCTGGCTTCTTACAATGATGAAAAAGACAAAGAGATGGTTTCTTCAATACTG 1788
Qy 1707 CTGCGATTGAGTCGGTGAATACTGCCAGCACTCTATGCAATGGTACTGATCT 1766
Db 1789 TTGTCCTAGATTGAGTCAGGACTGTCCAGTAATCTGCCATGGGATGGTGAATGTGT 1848

Qy 1767 CTGGGACCTGCCACTGCTTCTCTGGGTTTCTGGGCCCCGAGCTGTGGCAGAGCCTCTCTGCC 1826
Db 1849 CCGGGGTGTGTCACCTGTTTCCAGGATTTCTAGGAGCAGACTGTCTAAAGTGTCTCTGCC 1908
Qy 1827 CCGTGTCTCTGTAAGCGAAATGSCCAATACATGAAAGGAGATGCTTGTGCCACAGTGGCT 1886
Db 1909 CTGTCTGTGCACTGGGAATGGACAATATTCTAAAGGACGTCGCCAGTGTACAGCGCT 1968
Qy 1887 GGAAGGCGCTCAGTGCATGTGCCCAACACCACTGATCGATGTGGCTGTGCGCTGCAACAAC 1946
Db 1969 GGAAGGTCGAGAGTGCACGTCGCCATGAATCAGTGCATCGATCTCTTCTTCTGCGGGGCC 2028
Qy 1947 ATGGACCTGTGATCAGGCGCACCTCATCTGCAACCTTGGCTACAAGGCGCAGAGCTGTG 2006
Db 2029 ACGGCTCTGCAATGATGGGAATGTGTCTGTCTGTCTGCTACAAAGCGGAGCACTGTG 2088
Qy 2007 AGGAAGTGAATGCAATGAGACCCCAATGTTAGGCGGGGTGTCTGCTGAGAGCGCAAT 2066
Db 2089 AGGAAGTGAATGCTTGGATCCCACTGCTCAGCCACGAGTCTGTGTGAATGAGAAT 2148
Qy 2067 GCATTGCTTTGTGGATGGGAGGACCAACTGCGAGACCCCGAGGCCCACTGCTTAG 2126
Db 2149 GCTGTGAGCCCTGCTGGGTGTCTGAACTGTGAGCTGGCGAGGTCCAGTGCACCAG 2208
Qy 2127 ACCAGTGTTCAGGCCACCGAACTTCTTCCCGGACACCGGCTTTGCACTGTGACCCAA 2186
Db 2209 ACCAGTGCAGTGGGATGGCAGCTACCTGCTGACACCGGCTCTGCACTGCGATCCCA 2268
Qy 2187 GCTGACTGGACAGACTGTTCTATCGAGATCTGTGTCGCGACTGTGTTGGCCATGGCG 2246
Db 2269 ACTGATGGGTCCCAGCTGCTCTGTTGAAGTGTGCTCAGTAGACTGTGGCACTCAGCGCG 2328
Qy 2247 TGTGCTTGGGGGCACTTCCGCTCGAGGATGGCTGATGGGGGCGAGCTTGCAGCCAGC 2306
Db 2329 TCTGATCGGGGAGCTTCCGCTGTGAAGAGGCTGGACAGGCGAGCGTGTGACCAAGC 2388
Qy 2307 GGGCTTGCACCCCGCTGTGCGAGCATGGGACTGCGCGGAGCGGCAAGTGGAGTGCA 2366
Db 2389 GCGTGTGCCACCCCGCTGCAATGAGCATGGGACTGTAAGATGGCAAAATGTGAATGCC 2448
Qy 2367 GCGCTGGCTGAAATGGCGAACTGCACTGCACTCGCTCACTATCTGATAGGTTAGTAAAG 2426
Db 2449 GAGAGGCTGGAATGGTGAACACTGCACAT-----TG 2481
Qy 2427 AGGGTTGCTGGGTGTGCAATGGCAACCGCAGATGACCTTAGACTGTAATGGTTGGC 2486
Db 2482 ATGGCTGCCCTGACTTGTGCAACCGTAACCGGAGATGCACACTGGGTGAGAACAGCTGGC 2541
Qy 2487 ACTGCTGTGCGAGCTGGCTGGAGAGGAGCTGGCTGTGACACTTCCATGGAGACTGCT 2546
Db 2542 AGTGTGTCTGCGAGACCGCTGGAGAGGCGCCGAGTGCACGTTGCCATGGAATCTCT 2601
Qy 2547 GCGGTGACAGCAAGACAATGATGGAGATGGCTGTTGGACTGCAATGGACCCCTGACTGCT 2606
Db 2602 GTGCTGATAACAAGGATAATGAGGAGATGGCTGGTGGATTTGTTGGACCTGACTGCT 2661
Qy 2607 GCCTCAGCCCTGTGCCATATCAACCGCTGTGCTTGGCTTGGCTCCCTTAAACCTCTGACA 2666
Db 2662 GCCTCAGTCACTGTGTCAGAACAGCTGCTCTCCCGGGGTCCCGGAGCCCACTGGACA 2721
Qy 2667 TCATCCAGGACACAGGTCCCTGTGTGTCAGACAGAACCTACACTCTCTTATACCGCA 2726
Db 2722 TCATTCAG-----CAGGCCCAGAGGATTTGGCCCGAGTGAATGCTCTTATGACCGTA 2775
Qy 2727 TCAAGTTCTGTGGGAGGAGCAGCAGCAATAATCCCGGGAGAAACCCCTTTGATG 2786
Db 2776 TCAAGCTTTGGCAGGCAAGGATAGCACCCACATCTCTCTGGAGAGAACCCCTTTCAACA 2835
Qy 2787 GAGGCACTGTGTTGTTATTCGTGGCCAAAGTGAATGACATCAGATGGAACCCCTCTGTTG 2846
Db 2836 GCAGCTTGGTTTCTCTCATCCGAGGCCAAGTAGTAATCAAGATGGAATCTCCCTCTGTTG 2895

2847	QY	GTGTGAACATCAGTTTTGTCTCAATAAACCTCTCTTTTGGATATACAATCAGACGACCAAGTG	2906
2896	Db	GTGTGAACGTGTCTTTTGTCAAGTACCCAAAATACGGCTACACCATCACCCGCCAGGATG	2955
2907	QY	GCAGCTTTGACTTTGGTGACAAATGGGGCATCTCCATCATCTCCGGTTTCGAGCGGGCAC	2966
2956	Db	GCAGTTTCGACCTGATCGCAATGGAGGTGCTTCTTTGACTCTACACTTTGAGCGAGCCC	3015
2967	QY	CTTTTCATCACAGGAGGACACCCCTGTGGTGCATGGGATCGCTTCTTTGTTCATGGAAA	3026
3016	Db	CGTTTCATGACCCAGGAGCGCACTGTGTGGCTGCGGTGGAAACAGCTTTTAGCCCATGGACA	3075
3027	QY	CCATCATCATGACATGAGAGAAATGAGAAATCCCACTGTGACCTTGAGCAATTTTGGCCC	3086
3076	Db	CCCTGGTGATGAGACCGAGAGAACTTCATCCCGAGCTGTGACCTCAGTGGCTTTGTCTC	3135
3087	QY	GCOCCAACCAGTGTCTCTCCATCCCACTGACGTCCTTGGCGAGCTCCTGTGCAAGAGA	3146
3136	Db	GGCCTGATCAATCATCATCTCCTCCCACTGTCCACCTTCTTTAGTGTCTGCCCTGGGC	3195
3147	QY	AAGGCCCAATTGTGCCGGAAATTCAGAGCTTTGCAGAGGAAATCTCTATCTCTGGCTGCA	3206
3196	Db	AGAATCCCACTCGTGCCGTGAGACCCAGGTTCTTTATGAAGAAATCGAGCTCCCTTGGTTCCA	3255
3207	QY	AGATGAGGCTTGAGCTACTCTGAGCAGCGGACCCCTGGCTCAAAATCTGTCTCCTGAGATCA	3266
3256	Db	ATGTGAACCTTCGCTATCTGAGCTCTAGAACTGCAGGGTACAAGTCACTGCTGAAGATCA	3315
3267	QY	GCCTCACCACCGGACCATCCCTTCAAACCTCATGAAGGTGCACCTCATGGTACCGGTGG	3326
3316	Db	CCATGACCCCAAGTCCACAGTCCCCCTGAACCTCAATAGGTTTCACTGATGGTGGCTGTGG	3375
3327	QY	AGGGCCCTCTTTCAGGAAGTGGTTCCGCTGCAGCCCCAGACCTGTCCCTATTATTTCATTT	3386
3376	Db	AGGGGCATCTTTCAGAAAGTCATTCAGAGCTTCTCCCACTGGCTCCACCTTTCATCT	3435
3387	QY	GGGACAGAAGCAGAGTCTCAAAACGAAGGTGTTTGGGCTTTCAGAAAGCCTTTGTTTCCG	3446
3436	Db	GGGACAGAAGATGCGCTATGGCCAAAGGGTGATGACACTCAGATGCTGTTGTGTGTCG	3495
3447	QY	TGGGTTATGAATATGAATCTGCCACAGATCTAATCTGTGGGAAAAAGAACCAACAGTGC	3506
3496	Db	TCGGGTTTGAATATGAGACCTGTCCCAAGTCCTCCTAGTCCTAATCTCTGGGAAAGAGCAGCCCTCC	3555
3507	QY	TGCAGGCTATGAAATTTGACGCGTCCAAAGCTTGGAGGATGGAGCTTAGACAAACATCATG	3566
3556	Db	TTCAGGGATTCGAGCTGGACCCCTCCAACTCGGTGGTGGTCCCTTAGACAAACACCACA	3615
3567	QY	CCCTCACATTTCAAAGTGGTATCCTGCACAAAGGAATGGGGAAACCAAGTTTGTGTCTC	3626
3616	Db	TCCTCAATGTAAAGTGGAAATCCTACAAAGGCACTGGGGAAAAACAGTTTCTGACCC	3675
3627	QY	AGCAGCCTCTGTCTATTGGGAGCATCATGGGCAATGGGCGCCGGAGAAGCATCTCCTGGCC	3686
3676	Db	AGCAGCTGCCATCATCACCAGCATCATGGGCAATGTCGCGCCGGAGCATTTCCCTGTCTC	3735
3687	QY	CCAGCTCGAACCGGCTTCTGTGACGGGAAACAAAGCTCTCGGCCCAAGTGGCCCTCACCTGTG	3746
3736	Db	CCAGCTCGAACGGCTTGTCTGAAGGCAACCAAGCTGCTGGGCCCAAGTGGCTCTGGCTGTG	3795
3747	QY	GCTCTGACGGGAGCCTCTATGTGGTGATTTCAACTACTATTAGAAAGGATCTTCCCTCTCTG	3806
3796	Db	GAATCGATGGAGAGCCTCTATGTGGTGACTTCAATTTACATCCGAGCATCTTTCCTCTCTC	3855
3807	QY	GAATGTCAACCAACATCCTAGAGCTGAGGAATAAGATTTTCAGACATAGTCACAGTCCAG	3866
3856	Db	GAATGTGACCCAGCATCTTGAGGTTCAGAAATAAAGAGTTTAAACATAGCAACCAACCCAG	3915
3867	QY	CAGACAAATACTACTCTGCCACAGACCCCATGAGTGGGGCGCTCTTCTCTTCTGACAGCA	3926
3916	Db	CACACAAAGTACTACTTGGCAGTGAACCCCGTTCGGGCTCGCTCTACGTGTCCGACACCA	3975
3927	QY	ACAGCCGGCGGGTCTTTAAAATCAAGTCCCACTGTGTGGTGGTGAAGGACCTTGTCAAGAAGT	3986

3976	Db	ACAGCAGGAGAACTTACCGCGTCAAGTCTCTGAGTGGAAACCAAGACCTCGCTGGGAATT	4035
3987	Qy	CTGAGGTGGTTGCGGGACAGGTGAACAGTGCCTCCCTTTTGATGACACTCGCTGGGGG	4046
4036	Db	CGGAAGTTGTGGCAGGACGCGAGACAGTGTCTACCTTTTGATGAAGCCGCTGCGGG	4095
4047	Qy	ATGGTGGAGAGGCCACAGAGCCACACTCAACATCCACAGGGTATTACAGTGGACAAGT	4106
4096	Db	ATGAGGGGAAGGCCATAGATGCAACCTGATAGCCCGAGAGGTATTGCAGTAGACAGA	4155
4107	Qy	TTGGGCTGATCTACTTCGTGGATGGCACCATGATCAGACGCATCGATCAGAATGGGATCA	4166
4156	Db	ATGGGCTCATGTACTTTGTGATGCCACCATGATCGGAAGGTTGACCAATGGAATCA	4215
4167	Qy	TCTCCACCTGCTCGGCTCTAATGATCTACATCAGCCCGGCCACTCAGCTGTGATTCTG	4226
4216	Db	TCTCCACCTGCTGGGCTCCAAATGAOCTACTGCGCTCGGGCGCTGAGCTGTGATTCCA	4275
4227	Qy	TCATGGATATTTCCACGTAAGACTGGAGTGGCCACAGACTTAGCCCATCAACCCCAATGG	4286
4276	Db	GCATGGATAGCCCAGGTTCTGCTGGAGTGGCCACAGACCTTGTGTCAATCCCATGG	4335
4287	Qy	ACAACACTCATTTATGTCTCGACAAACATGTGTGTCGTGAAATCTCTGAAAAACCAACAGG	4346
4336	Db	ATAACTCTTTGATGTTCTAGAGAACAAATGTCACTCTTCGAATCACCGAAGACCAACAG	4395
4347	Qy	TGCGCATGTGCGCGGAGGCCCATGCACATGCCAGGTCCCTGGCATTAGCACCTTCTCTGC	4406
4396	Db	TCAGATCATTTGCGGAGCGCCCATGCATCTGCCAAGTTCTTGGCATTTGACTACT--CAC	4452
4407	Qy	TAAACAAGGTGGCCATCACGCAACCTTGGAGTCAGCCACCGTTTGGCTGTTTTCACACA	4466
4453	Db	TCAGCAACTAGCCATTCACTCTGCCCTCGAGTCAGCCAGTCGCATTTGCCATTTCTCACA	4512
4467	Qy	ATGGGTCTCGTATATTGCTGAGACTCATGAGAAAGATCAACCGCATCAGGCAGCTCA	4526
4513	Db	CTGGGTCTCTATCATCACTGAGACATGAGAGAAAGATTAAACCGTCTAGCCACGGTAA	4572
4527	Qy	CCACTAGTGGAGAGATCTCACTCTGCTGGGGCCCCAGTGGGTGTGACTGTAAAAATG	4586
4573	Db	CAACCAACGGGAGATCTGCCTTTTACCTGGGCAGCCTCGGACTGCGACTGCACAAAACG	4632
4587	Qy	ATGCAACTGTGATGTTTTTCTGGAGACATGTTATGCAAGGATGCAAAAGTTAAATA	4646
4633	Db	ATGTCAATTGCACTGTCTATTACGAGAGATGATGCTACGCGACTGATGCCATCTTGAATT	4692
4647	Qy	CCCCATCTTCCTTCGCTGTGTGCTCATGGGAGCTTACGTGGCGCGACCTTCGGGAACA	4706
4693	Db	CCCCATCATCTTACCTTAGCTCCAGATGGTACCATTTCATATGACAGCCTTGGAAATA	4752
4707	Qy	TCCGAATTCCGTTTATCCGGAAGAACAGGCTTTCTCAACACCCAGAACATGTATGAGC	4766
4753	Db	TTCCGATCAGGCGGCTCAGCAAGAACAGGCTGTTCTTAATGCTTCAACCCAGTATGAGG	4812
4767	Qy	TGTCCTTACCAATTTGACACGAGCTCTATCTGTTTGATACACCCGCAAGCAGCTGTACA	4826
4813	Db	CTGCATCCCCCGGAGACGAGGTTATGTTTTTCAACGCTGATGGCATCCACCAATTACA	4872
4827	Qy	CCCAAGCCTGCCACACGAGACTACCTGTACAACTTCCACTACATCGGGACGGCGACA	4886
4873	Db	CTGTGCGCTGTGACAGGGGAGTACTTGTACAAATTTACATAAGTACTGACAATGATG	4932
4887	Qy	TCACACTCATCACAGACAAACATGGCAACATGTTAAATGTCCGCGGAGACTCTACTGGGA	4946
4933	Db	TCACTGAATTGAATTAATGGGAAATTCCTTGAAGATCCGTCGGGACAGCAGTGGCA	4992
4947	Qy	TGCCCTCTGCGTGGTCCAGATGGCCAGGTGTACTGGGTCAACCATGGGCACCAACA	5006
4993	Db	TGCCCGCTCACCTGCTATGCTTGACAAACGATCATCAACCTCACCGTGGGCACCAATG	5052
5007	Qy	GTGCATCTCAAGAGTGTGACCACAACAGACACAGAGTTGGCCCATGATGACATACCATGGCA	5066

Db 5053 GAGGCTCAAGTCTGTCTCCACACAGAACCTGGAGCTTGGTCTCATGACCTATGATGGCA 5112
 QY 5067 AFTCCGGCTTCTGGCAACCAAGCAATGAAGCAATGAAGCAACATTTTATGATGACG 5126
 Db 5113 ACATCGGCTCTCTGGCCACCAAGACGATGAACAGAGTGAACGACTTTCTATGACTATG 5172
 QY 5127 ACAGCTTTGGCGCTGACCAATGTACCTTCCCTACTGCGGACGAGTGAAGTTCCGAA 5186
 Db 5173 ACCAGAGGCGCTGACCAAGTGAACGCGCCCAACGGGGTGAACCAAGTCTGCACC 5232
 QY 5187 GTGATACAGACAGTTTCACTGATGCTCCAGGTAGAGACCTCCA---GCAAGGATGATGCA 5243
 Db 5233 GGGAAATGGAGAAATCTATTACCAATGACATTGAGAACTCCAAACCGTGATGATGACGTCA 5292
 QY 5244 CCATAACCAACCACTGTCTGCTCAGCGCTTCTACACACTGCTGCAAGACCAAGTCC 5303
 Db 5293 CTGTCAATCACCACACCTCTCTTCACTAGAGGCTCTCTACACAGTGGTACAAGATCAAGTTC 5352
 QY 5304 GGAACAGCTACTTACATCGGGGCGGATGGCTCTTGGCGCTGCTGCTGCGCCAAACGGCATGG 5363
 Db 5353 GGAACAGCTACCAGCTCTGTAAATGTAACCTGAGGGTATGATGCTAAATGGGATGG 5412
 QY 5364 AGTGGCGCTGACAGTACTGAGCCCACTTGTGCTGAGCAACGTCGCAACCCACCGTGGGCA 5423
 Db 5413 GTATCAGCTTCCACAGCAGGCGCCATGTCTTAGCGGGCACCATCACCCCCACCATTTGGAC 5472
 QY 5424 AGAGAAATGTCAAGCTGCGCCATCGACCAACGGCTCAACCTGGTGGAGTGGCGGCGAGGCA 5483
 Db 5473 GCTGCACATCTCCCTGCTATGAGAAATGGCTTAACTCCATTGAGTGGCGCTAAGAA 5532
 QY 5484 AAGACAGCTCGGGGCGAGTCACTGTCTTTGGGCGCGGCTGCGGGTGCACAAACCGAA 5543
 Db 5533 AGGAACAGATTAAGAGGCAAGTCAACCTCTTTGGCAGGAAGTCCGGGCTCATGGAAGAA 5592
 QY 5544 ATCTCCTATCTCTGACTTTCATCGCTGACACACGACAGAGAGATCTATGATGACCAAC 5603
 Db 5593 ATCTCTGCTCAATGACTATGATGAAATATTCGGACTGAAAGATCTATGATGACCAAC 5652
 QY 5604 GCAAGTTCACCTTTCCGATCTGTACGACAGCGGGGCGGCGCCAGCTCTGTGTCAACCA 5663
 Db 5653 GGAAGTTCAACCTGAGGATCATTTATGACAGGTGGCGCGCCCTTCTCTGCTGCGCCA 5712
 QY 5664 GCAGAGCTGAATGGTGTCAACGTGACATCTCCCTGGGGGTATCATTTGCTGGATCC 5723
 Db 5713 GCAGCGGCTGGCAGCTGTCAACGTGTCACTTTCAATGGGCGCTGCTGGGCTTC 5772
 QY 5724 AGAGGGGCATCATGTCTGAAGAATGAATACGACAGCGGGCGCGGATCACATCCAGGA 5783
 Db 5773 AGCGTGGGCGCATGACGAGAGGACACACATCGACAGCAAGCGCGATCGTGTCCGCA 5832
 QY 5784 TCTTCGCTGATGGGAAGACATGGAGCTTACATCTTTAGAGAGTTCATCGTCTGCTAC 5843
 Db 5833 TGTTCGCTGACGGGAAAGTGTGGAGCTACTCTACCTTGAACAGTCCATGGTCTCTCTGC 5892
 QY 5844 TACACAGCCAGAGCAGTATATCTTTGAGTTCGACAGAGATGACCGCTCTCTTCTGTGA 5903
 Db 5893 TTCAGAGCCAAAGTCAAGTATATTTGAGTATGACTCTCTGACCGCTCTCTTGGCGTCA 5952
 QY 5904 CGATGCGCAACGTTGGCGGCGGACACACTAGAGACCAATCCCGCTCAGTGGGTACTACAGAA 5963
 Db 5953 CCATGCGCAGGCTGGCGCGGACAGCATGTCCACACACACTCCATCGGCTACATCCGTA 6012
 QY 5964 ACATCTATCCCGCTTGGGCAATGCGCTCAGTGCATACAGGACTTCACTGAGAGTGGC 6023
 Db 6013 ATATTTCAACCCGCTGAAAGCAATGCTTCGGTCACTTTGACTACAGTGAAGCGGC 6072
 QY 6024 ACCTCCTTCAACCTTCTACCTGGGCACTGGCGGAGGTGATATCAAGTATGGCAAAAC 6083
 Db 6073 GCATCCTGAAGACCTCTTTTGGGCAACCGGACCGCAGGTGTTCTCAAGTATGGGAAC 6132
 QY 6084 TGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGTCAAGTTCACCTATGACGAGA 6143
 Db 6133 TCTCCAAGTTATCAGAGATTGTCTACGACAGTACCGCGCTCACCTTGGGTATGACGAGA 6192

QY 6144 CGGCGGCGATGCTGAAGACCATCAACCTACAGATGAGGGCTTCACTTGCACCATCCGCT 6203
 Db 6193 CCACCTGGTGTCTTGAAGATGGTCAACCTCCAAAGTGGGGCTTCTCTTGCACCATCAGT 6252
 QY 6204 ACCGTGAGATTGGGCGCTGATTGACCGACAGATCTTCCGCTTCACTGAGGAAGCATGG 6263
 Db 6253 ACCGGAAGATTGGCCCCCTGGTGGCAAGCAGATCTACAGGTTCTCCGAGGAGCATGG 6312
 QY 6264 TCAACGCCGTTTTGACTACAACT--ATGACAAACAGCTTCCGGGTGACCAAGATGACGG 6320
 Db 6313 TCAATGCCAGGTTTGACTACATCACTATGACCAACAGCTTCCGATCGCAAGCATCAAGC 6372
 QY 6321 CTGTGATCAACAGACCCCACTGCGCATTTGATCTCTATCGCTATGATGATGTGTCAGGCA 6380
 Db 6373 CCGTCATAGTGAGACTCTCCCTCCCGCTGACCTCTACCGCTATGATGAGATTTCTGGCA 6432
 QY 6381 AGACAGACAGTTTGGGAAGTTTGGTGTCACTTATGATGACATTAACCAAGATCATCAACA 6440
 Db 6433 AGGTGGAACACACTTGTGAAGTTTGGAGTCACTATTAATGACATCAACCAAGATCATCAACA 6492
 QY 6441 CAGCTGTCAATGACCCACACCAAGCATTTTGTATGATATGGCAGGATGAAGAGTGCAGT 6500
 Db 6493 CTGCGCTGATGACCTCAGCAACACTTTCGACACCCATGGGCGGATCAAGGAGTCCAGT 6552
 QY 6501 ATGAGATCTTCGCTCGCTCATGTACTGATGACCGTCCAGTATGATTAACATGCGGCGAG 6560
 Db 6553 ATGAGATCTTCGCTCGCTCATGTACTGATGACCGTGCATATGACAGCATGGCGAGG 6612
 QY 6561 TAGTGAAGAGAGCTGAAGGTAGGACCTTACGCGCAATACCACTGCTACTCTTATGAGT 6620
 Db 6613 TGATCAAGAGGAGCTAAACCTGGGCGCTTATGCGCAATACCAAGAGTACCACTATGACT 6672
 QY 6621 ATGATGCTGACGGCGAGCTGACAGCAGTCTCATCAATGACAGCACTCTGCGGCTACA 6680
 Db 6673 AGATGGGACGGGCGAGTCCAGAGCGTGGCGTCAATGACCGCGACCTTGGGCTTACA 6732
 QY 6681 GTACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGGGAAACAGTGCAGGCTCACAC 6740
 Db 6733 GCTATGACTTAAATGGGAATCTCCACTTACTGAAACCCAGGCAACAGTGTGCGCTCATGC 6792
 QY 6741 CACTACGCTATGACATCGCGACCGCATCACTCGGCTGGTGAAGTCAATACAGATGG 6800
 Db 6793 CTTGCGTATGACCTCGGATCGGATACAGACTCGGGATGTGCGATACAAATG 6852
 QY 6801 ATGAGGATGGTTCCTGAGGCGAGGGGGGTGATATCTTTGAGTACAACTCAGCTGGCC 6860
 Db 6853 AGGACGATGGCTATCTGTGCCAGAGAGGGTCTGACATCTTTCGAATACAAATCCAGGGCC 6912
 QY 6861 TGCTCATCAAGGCTTAAACCGGGCTGGCAGCTGAGTGTGAGTACCGCTACGATGGCC 6920
 Db 6913 TCCTAAAGAGGCTTAAACCAAGGCGCAGCGGTGAGTGTCCAGTACCGCTATGATGGCG 6972
 QY 6921 TGGGCGCGGCTGTCCAGCAAGAGCAGCCACAGCCACTGCACTCTTCTTATGAG 6980
 Db 6973 TAGGACGGGGCTTCTTACAGAACCACTTGGGCGCCACTGAGTACTTCTACTCTG 7032
 QY 6981 ACCTGACCAACCCCAAGGTACCCACTGTATCAACCACTCAGCTGTGAGATCACT 7040
 Db 7033 ACCTCAACACCGGCGCATCACTCATGCTCACTCACTCTCGGAGATTACCT 7092
 QY 7041 CCCTCTACTAGCTTGAAGGACACCTTTTGGCATGGAGCTGAGGATGATGATGAGT 7100
 Db 7093 CACTGTACTAGCTTCCAGGGCCACTCTTTTGGCATGGAGCAGGATGGGAGGAGT 7152
 QY 7101 TTTACATAGCTTGTGACAACTCGGGACCCCTCTTGTGTCTTTAGTGGAAACAGGTTTGA 7160
 Db 7153 ACTATGTTGCTCTGTATTAACACAGGACTCTCTCTGGCTGTGTTTCAAGATCAACGGCTCA 7212
 QY 7161 TGATCAAGCAATCTGTACAGAGCTTATGGGAGATCTACATGATACCAACCCCACT 7220
 Db 7213 TGATCAACAGCTGAGTACACAGGCTTATGGGAGATTATTTATGACTCCAAACCCGACT 7272

QY 7221 TTCAGATCATCATAGGCTACATGGTGGCTCTATGATCCACTCACCAGCTGTGTCACA 7280
 Db 7273 TCCAGATGGCTATTTGGCTTCATGGGGAGCTCTATGACCCCTGACCAAGCTGGTCCACT 7332
 QY 7281 TGGCCCGCGAGATTATGATGTGCTGCGGAGCTGAGTACGCCAGACCAAGAGCTGT 7340
 Db 7333 TCATCAGCTGATATGATGTGCTGCGAGAGATGAGTACCTCCAGACTATACCATGT 7392
 QY 7341 GGAAGCACCTTATGAGCAGCAAGTCAATGCTTTTAACTCTATATGTTTCAAAAAACA 7400
 Db 7393 GGAAGAACCTT---GGGCAAGAGAGCGCGCCCTTTAACTCTATATGTTTCAAGAGCAACA 7449
 QY 7401 ACCCATCATCAAGCTCCAGGACATCAAGTGTCTCATGACAGATGTTTAAAGCTGGCTGC 7460
 Db 7450 ATCTCTCAGCAGTGAAGTATGTTGAGAACTACGTGACAGATGTAAGAGCTGGCTTG 7509
 QY 7461 TCACCTTTGGATTCAGCTACACAGCTGATCCCTGTTTATCCCAACACAGATGATG 7520
 Db 7510 TGATGTTTGGATTCAGCTTAGCAACATCATCTCTGCTTCCCGAGAGCCAAATGTATT 7569
 QY 7521 CCATGGACCTCTCAGAGCTCATCCACACAGATGAAACGCGAGAGTGGCAACA 7580
 Db 7570 TCGTGCCTCTCCCTATGA-----ATTGTCAGAGAGTCAAGCAAGTGAAGATG 7617
 QY 7581 GCAAGTCTATCTCGGGGTACAGTGTGAAGTACAGAGAGCTCAAGCCCTTTGTCACT 7640
 Db 7618 GACAGCTATTACAGGTGTCCACAGACACAGAGACATACACAGCCCTTCATGGCTC 7677
 QY 7641 TAGAAGCTTTGACAGCTTATGGCTTCCCAATCACAGCTGCCAGAGCTCCCAAGA 7700
 Db 7678 TGGAA-----GGACAGCTATTACTATAAAGCTCCACGCGCATCCGAGAGAAAGCAG 7731
 QY 7701 CCAAGAGTTTGCATCCAGCGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAG 7760
 Db 7732 GTCAGTGTTCGCCACCAACAGCCCATCATTTGGCAAGGATCATGTTTGCATCAAG 7791
 QY 7761 ATGGCCGAGTACACACAGATCATCATGTTGGTCAATGAGATGGCGCAAGGTTGCTG 7820
 Db 7792 AAGGCGGGTCAACAGCGGTGTCCAGCATCGCAGCAAGATAGCGCAAGGTGGCAT 7851
 QY 7821 CCATCTTGAACCATGCCACTACTAGAGAACCTGCACTTCACTATGATGGGTGATA 7880
 Db 7852 CTGTGCTGAACAGCGCTACTACTGTGCAAGATGCACTACAGCATCGAGGGCAAGGACA 7911
 QY 7881 CCCATTACTTTGTGAACCCAGGACCTTCAGAGGTGACCTGGCCATCTCTGGGCTCAGT 7940
 Db 7912 CCCACTACTTTGTGAAGTTGGCTCAGCCGATGCGACCTGGTCACTAGGACACCA 7971
 QY 7941 GGGGCGGCGAAACCTGTAGAGATGGGTCAACGTCACTGTGTCCAGATCAACACAGTAC 8000
 Db 7972 TCGGCGCGCAAGGTGTAGAGAGCGGGGTGAACGTGAACCGTGTCCAGCCACCGTGTGG 8031
 QY 8001 TTAATGGCAGACTAGACGCTACACAGACATCCAGCTCCAGTACGGGCACTGTGCTGA 8060
 Db 8032 TCAAGCGGAGGACTCGAAGTTACGAAATGAGTTCAGATCTCCAGTCTCCAGCTGTGCTCA 8091
 QY 8061 ACACACGCTACGG-----GACAGGTTGGATGAGGAGAGGACGCGTCTCTGGAGC 8111
 Db 8092 GCATCCGCTATGGCTCATCCCGCCGACACCTTGGACGAGAGAGGCGCGGTCTCTGGACC 8151
 QY 8112 TGGCCCGCAGAGACCGCTGCGCAAGCTGGGCGCGGAGCAGCAGAGACTCGCGAAG 8171
 Db 8152 AGGCGAGACAGAGGCGCTGGGCAAGCTGGGCGCGGAGCAGCAGAGAGGCGGAGCG 8211
 QY 8172 GGGAGGAAGGCTCGCGGCTTGGACAGAGGGGGGAGAGCAGAGGTGCTGAGCAGAGCGG 8231
 Db 8212 GGAGAGGGGAGCGCGCTGTGGACTGAGGGCGAGAGAGCAGCAGCTTCTGAGCAGCGGG 8271
 QY 8232 GGGTGCAAGGTACAGCGCTTTTTCGTGATCTCTGTGAGCAGTACCCAGAACTGTGAG 8291
 Db 8272 GCGTGCAAGGTTACAGGGATATTACGTGCTTCCCGTGGAGCAATACCCAGAGCTTCAG 8331
 QY 8292 ACAGCGCCCAACAATCCATTCATGACAGAGGAGATGGCGCGGAGGTGACAGA 8348

Db 8332 ACAGTAGCAGCAACATCCAGTTTAAAGACAGATGAGATGGAAAGAGGTAAACAA 8389
 RESULT 11
 ADB32028
 ID ADB32028 standard; cDNA; 9729 BP.
 XX
 AC ADB32028;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human FCTR3f cDNA.
 XX
 KW Human; gens; ss; FCTR3f; colorectal cancer; adenomatous polyposis coli;
 KW myelogenous leukemia; congenital neonatal autoimmune thrombocytopenia;
 KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
 KW renal cell carcinoma; melanoma; clear cell carcinoma;
 KW granular cell carcinoma; neurological disorder;
 KW neurodegenerative disorder; nerve trauma;
 KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
 KW Gardner syndrome; mental health condition; immunological disorder;
 KW allergy; asthma; lung disease; reproductive disorder; deafness;
 KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
 KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
 KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
 KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
 KW lattice corneal dystrophy.
 XX
 OS Homo sapiens.
 XX
 PN US2003087816-A1.
 XX
 PD 08-MAY-2003.
 XX
 PF 05-MAR-2001; 2001US-00800199.
 XX
 PR 03-MAR-2000; 2000US-0186592P.
 XX
 PA (VERM/) VERMET C.
 PA (FERN/) FERNANDES E.
 PA (SHIM/) SHIMKETS R.
 PA (HERR/) HERRMANN J.
 PA (MAJU/) MAJUMDER K.
 PA (MACD/) MACDOUGALL J.
 PA (MISH/) MISHRA V.
 PA (MEZE/) MEZES P S.
 PA (RAST/) RASTELLI L.
 XX
 PI Vermet C, Fernandes E, Shimkets R, Herrmann J, Majumder K,
 PI Macdougall J, Mishra V, Mezes PS, Rastelli L;
 XX
 WP1: 2003-525633/59.
 P-PSDB: ADB32029.
 XX
 PT New FCTRX polypeptide and encoding polynucleotide, useful for preventing
 PT or treating FCTRX-related disorders, such as cancer, autoimmune,
 PT neurodegenerative, gastrointestinal, reproductive and inflammatory
 PT diseases.
 XX
 PS Claim 9; Page 34-37; 155pp; English.
 XX
 CC The invention relates to FCTRX polypeptides and the polynucleotides
 CC encoding them. The sequences of the invention are useful for the
 CC manufacture of a medicament for diagnosing and treating disorders
 CC associated with the FCTRX polypeptide, such as colorectal cancer,
 CC adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
 CC autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
 CC tumours, mammary tumours, human gliomas, astrocytomas, renal cell
 CC carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
 CC cell and granular cell carcinomas, neurological disorders,
 CC neurodegenerative disorders, nerve trauma, familial myelodysplastic
 CC syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health

QY 2007 AGGAAGTGGAGTGCATGAGACCCACATGTTTCAGGCCGCGGTGTCTCGTGCAGAGCGGAAT 2066
Db 2089 AGGAAGTGAATGCTGGATCCCACTGCTCCAGCCACGGAGTCTGTGTGAATGGAGAAT 2148
QY 2067 GCCATTGCTTTGTGGATGGGAGGACCACTGCGAGACCCCGAGGCCACACATGCTTAG 2126
Db 2149 GCCTGTGAGCCCTTGGCTGGGGTGGTCTGAATGTGAGCTGGCGAGGGTCCAGTGGCCAG 2208
QY 2127 ACCAGTGTTCAGGCCACCGGAACCTTCTCCCGGACACCGGGCTTTGAGAGTGTGACCCAA 2186
Db 2209 ACCAGTGCAGTGGCATGGCAGCTACCTGCTGCACACGGGCTCTGCAGCTGCGATCCCA 2268
QY 2187 GCTGGACTGGACACAGACTGTTCTATCAGATCTGTGCTGCGGACTGTGTTGGCGATGGCG 2246
Db 2269 ACTGGATGGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2328
QY 2247 TGTGCTAGGGGGACCTGCGCTGCGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2306
Db 2329 TCTGCATCGGGGAGCTGCGCTGTGAGAGGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAG 2388
QY 2307 GGGCTTGCACCCGCTGTGCGAGCATGGGACCTGCGGCGAGCGCAAGTGGAGTGCA 2366
Db 2389 GCGTGTGCAACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2448
QY 2367 GCCCTGGTGAATGGGGAACACATGCACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2426
Db 2449 GAGAGGCTGGAATGGTGAACACTGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2481
QY 2427 AGGCTTGCCTGGTGTGCAATGGCAACGCGAGATGTACTTACAGCTGCAATGGTGGC 2486
Db 2482 ATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2541
QY 2487 ACTG 2546
Db 2542 AGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2601
QY 2547 GCGGTGACAGCAAGCAATGATGAGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2606
Db 2602 GTGCTGATAACAAGATGATGAGGAGATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2661
QY 2607 GCCTCCAGCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2666
Db 2662 GCCTG 2721
QY 2667 TCATCCAGGACACAGTCCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2726
Db 2722 TCATTCAG-----CAGGCCAGAGGATGGCCCGCAGTGAAGTCTTCTATGACCGTA 2775
QY 2727 TCAAGTTCCTGCTGGGAGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2786
Db 2776 TCAAGCTCTTGGCAGGCAAGGATAGCACCACATCATTCTTGGAGAGAACCTTTCAACA 2835
QY 2787 GAGGCGATGCTGTGTTATGCTGGCCAGTGTAGATCAGATGGGAACCCCTGCTGCTG 2846
Db 2836 GCAGCTTGGTCTCTATCCGAGGCCAAGTAGTAACTACAGATGGGAACCTCCCTGCTGCTG 2895
QY 2847 GTGTGAACATCAGTTTGTCAATAACCTCTCTTGTGATATACAACTCAGCAGGCAAGATG 2906
Db 2896 GTGTGAACGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2955
QY 2907 GCAGCTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2966
Db 2956 GCAGCTTGAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3015
QY 2967 CTTTTCATCACAGGAGCACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3026
Db 3016 CGTTTCATGAGCAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3075
QY 3027 CCATCATCATGAGATGAGAGATGAGATTCACGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3086
Db 3076 CCCTGGTGTGAAGACCGAGGAGAACTCCATCCCGCAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 3135
QY 3087 GCCCCAAACCCAGTGTCTCTCCATCCCACTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3146

Db 3136 GGCCTGATCCAATCATCTCTCTCCCACTGCTCCCACTCTCTTGTAGTGTGCTCCCTTGGGC 3195
QY 3147 AAGGCCCATCTGTCGCGAATTCAGGCTTTCAGGAGGAAATCTCTATCTCTGCTGCA 3206
Db 3196 AGAATCCCATCTGCTGCTGAGACCCAGTTCCTTATGAAGAAATCAGACTCCCTGTTTCCA 3255
QY 3207 AGATGAGGCTGAGTACCTGAGCAGCAGCCGACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3266
Db 3256 ATGTGAACCTTCGCTATCTGAGCTCTAGAACTGCAAGGTACAACTCACTGCTGAAGATCA 3315
QY 3267 GCCTCACCACCCGACCATCCCTTCAACCTCATGAAGTGCACCTCATGTTAGCGGTG 3326
Db 3316 CCATGACCCAGTCCACAGTGGCCCTGAACCTCATTAGGTTTCCCTGATGGTGGTGTG 3375
QY 3327 AAGGCCCTCTTTCAGGAAGTGGTTCGCTGACGCCCGACGCTGCTCTCTATTTATTTCAITTT 3386
Db 3376 AAGGGCATCTCTCCAGAACTCATTCAGGCTTCTCCCAACCTGCTCCACCTCATCT 3435
QY 3387 GGGCAAGACAGAGCTCTAACACAGAGGTTGTTGGGCTTTCAGAGCTTGTGTTGCG 3446
Db 3436 GGGCAAGACAGAGCTGATGGCCAAAGGGTGTATGGACTCTCAGATGCTGTGTTGCTG 3495
QY 3447 TGGGTTATGAATATGAATCCTGCCAGATCTAATCTGTGGGAAAGAAACAAAGTGC 3506
Db 3496 TCGGTTTGAATATGAGACCTGCTCCAGTCTTAATCTCTGGGAAAGGACAGCCCTCC 3555
QY 3507 TCAGGGCTATGAATGACGGTCCAAGCTTGGAGGATGGAGCTTAGACAAACATCATG 3566
Db 3556 TTCAAGGATTCAGGCTGACCCCTCCAACCTCGTGGTGTGCTCTAGACAAACACCA 3615
QY 3567 CCTCAACATTCAGAGTGGTATCTGTCACAAAGGAAATGGGAGAACAGTTTGTGCTC 3626
Db 3616 TCTCAATGTTAAAGTGGAACTCTACAAAGGCACTGGGAAACAGTTTCTTGACCC 3675
QY 3627 AGCAGCTCTCTGCTATGGGAGCATCATGGGCAATGGGCGCGGAGAGCACTCTCCTGCC 3686
Db 3676 AGCAGCTCTGCTCATCACAGCATCATGGGCAATGGTTCGCGCGGAGCAITTTCTGTC 3735
QY 3687 CGAGCTGAAACCGCTTGTGTCGCGCAACAGCTCTGCGCCAGTGGCCCTCATCTGTC 3746
Db 3736 CGAGCTGAAACCGCTTGTGTCGCGCAACAGCTCTGCGCCAGTGGCTCTGCTGCTGTC 3795
QY 3747 GCTCTGACGGGAGCTCTATGTTGGTGTATTCACATACATTAAGAGGATCTTCCCTCTG 3806
Db 3796 GAATCGATGGAGCTCTATGTTGGTGTACTCAATTAATCCGAGCATCTTTCCTCTC 3855
QY 3807 GAAATGTCAACCAATCTTAGAGCTGAGGAATAAGATTTACAGATAGTCAAGTCCAG 3866
Db 3856 GAAATGTGACAGCATCTTGGAGTTACGAAATAAGAGTTTAAACATAGCAACACCCAG 3915
QY 3867 CACACAAATACCTACCTGGCCACAGACCCCATGAGTGGGCGCTCTCTTCTGACAGCA 3926
Db 3916 CACACAAATACCTACCTGGCAGTGGACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3975
QY 3927 ACAGCGCGGCTCTTAAATCAAGTCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3986
Db 3976 ACAGCAGGAGATCTACCGCTCAAGTCTCTGAGTGGAAACCAAGACCTGGCTGGAAAT 4035
QY 3987 CTGAGTGTGCTGGGAGCAGGTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4046
Db 4036 CGGAAGTGTGCGAGGACGGAGAGCAGTGTCTACCTTTGATGAAGCCCTGCGGG 4095
QY 4047 ATGTGGGAAGGCCACAGAAAGCCACATCACCAAATCCAGAGGTTATACAGTGGCAAGT 4106
Db 4096 ATGGAGGGAAGGCCATAGATGCAACCTGATGAGCCGAGAGGATTTGCAGTAGACA 4155
QY 4107 TTGGGCTGATCTACTGCTGGATGCAACATGATCAGAGCATTCGATCAGAAATGGGATCA 4166
Db 4156 ATGGGCTGATCTACTGCTGGATGCAACATGATCAGAGGTTTACCAAGATGGAATCA 4215
QY 4167 TCTCCACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4226

Db 4216 TCTCACCCTCTGGCTCAATGACCTCACTGCGCTGCGGCGCTGAGCTGTGATCCA 4275
 Qy 4227 TCATGGATATTTCCAGTAAGACTGGAGTGGCCACAGACTTTAGCCATCAACCAATGG 4286
 Db 4276 GCATGGATGAGCCAGGTTCTGTGGAGTGGCCAAACAGACCTTGTGTCAATCCCATGG 4335
 Qy 4287 ACAACTCACTTTATGTCTCGACAAATGTGTCTCGAAATCTCTGAAACACCAAGG 4346
 Db 4336 ATAACTCCTGTATGTTCTAGAGAACAAATGTCATCTTCGAATCACCAGAACCAACAAG 4395
 Qy 4347 TGGCAATGTCCCGGAGGCCATGCACTGCCAGTCCCTGGCAATGACCACTTCTGTC 4406
 Db 4396 TCAGCATATTTGCGGACGCCCATGCACTGCCAAGTTCCTGCAATGACTACT--CAC 4452
 Qy 4407 TAAGAAAGTGGCCATCCAGCAACCTCTGGAGTCAGCCACCGCTTTGGCTGTGTTCAACA 4466
 Db 4453 TCAGCAAACTAGCCATTCACCTCTGCGCTGGAGTCAGCCAGTGCATTTCTCACA 4512
 Qy 4467 ATGGGCTCCTGTATATGCTGAGACTGATGAGAAAAGATCAACCGCATCAGGCAGTCA 4526
 Db 4513 CTGGGTCCTTACATCACTGAGACAGATGAGAAAGATTAACCGTCTAGCCAGGTAA 4572
 Qy 4527 CCATAGTGGAGAGATCTCACTCGTCTGGGCCCCCGCTGAGTGTGTAATAATG 4586
 Db 4573 CAACCAACGGGAGATCTGCTTTAGCTGGGCGAGCTCGGACTGCGACTGCAAAAACG 4632
 Qy 4587 ATGCCAAGTGTGATGTTTCTGGAGACGATGTTATGCCAAGGATGCAAGTTAAATA 4646
 Db 4633 ATGTCAATGCACTGCTATTCAGAGATGATGCTACGGAGTGTGCTCATCTTGAAT 4692
 Qy 4647 CCCCATCTCTGCTGTGTCTGATGGGAGCTCTACGTGGCGGACCTTGGGAACA 4706
 Db 4693 CCCCATCATCTTAGCTGAGTCCAGATGGTACCATTTACATTCAGACCTTTGGAATA 4752
 Qy 4707 TCCGAATCGGTTATTCGGNAGAACAGGCTTTTCCTCAACACCCAGAACATGTATGAGC 4766
 Db 4753 TTCGGATCAGGGCGGTGAGCAAGAACAGGCTGTCTTAATGCTTCAACCCAGTATGAG 4812
 Qy 4767 TGTCTTCAACAAATGACAGAGCTCTATCTGTTGATACACCGGCAAGCACCTGTACA 4826
 Db 4813 CTGATCCCCGGAGAGAGAGTTATGTTTCAACGCTGATGCAATCCCAATACA 4872
 Qy 4827 CCCAAAGCTGCCACAGGAGACTACCTGTACAACTTCACTACACTGGGAGCGGCA 4886
 Db 4873 CTGTGAGCTGTGACAGGGGAGTACTGTGACAAATTTCAATATAGTACTGACAAATG 4932
 Qy 4887 TCACACTCATACAGACAAATGGCAACATGGTAAATGTCGCCGAGACTTACTGGGA 4946
 Db 4933 TCATCGAATTTGACAAATGGGAATTCCTGAAAGATCCGTGGGAGCAGGTGGCA 4992
 Qy 4947 TGGCCCTCTGGCTGGTCCCAGATGGCCAGGTGTACTGGGTGACCATGGGCAACCA 5006
 Db 4993 TGGCCCTGCTCTCATGCTGACAAACAGATCATCACTCCCTCAACCGTGGCAACCAATG 5052
 Qy 5007 GTGCACCTCAGAGTGTACACACAGGACACAGATGGCCATGATGACATACCATGGCA 5066
 Db 5053 GAGCCCTCAAAAGTCTGTCCACACAGAACCTGGAGCTTGTCTCATGACCTATGATGGCA 5112
 Qy 5067 ATTCGGGCTTCTGGCAACAAAGCAATGAAACGGATGGACAAATTTATGATAG 5126
 Db 5113 ACATGGGCTCTGGCCACCAAGCGATGAACAGGATGGACGACTTCTATGACTATG 5172
 Qy 5127 ACAGCTTGGCGCTGACAAATGTGACCTTCCCTACTGCGCAGGTGAGCAGTTCCGAA 5186
 Db 5173 ACCAGAGGCGCTCACCAAGTGAAGGAAATTCCTGAAAGATCCGTGGGAGCAGGTGGCA 5232
 Qy 5187 GTGATACAGACAGTTCACTGATGCTCCAGGTAGAGACTTCCA--GCAAGGATGATGCA 5243
 Db 5233 GGGAAATGGGAATCTATTACCATTCACATTTGAGAACTCCACCGTGTGATGATGACTCA 5292
 Qy 5244 CCATAACCAACCTCTGCTCAGCGGCTTCTACACACTGCTGCAAGACCAAGTCC 5303
 Db 5293 CTGTATACCAACCTCTCTTCTAGTAGAGGCTTCTTACACAGTGGTACAAGATCAAGTTT 5352

Qy 5304 GGAAACAGCTACTACATCGGGCCGATGGCTCTTGGGCTCTGCTGGCCAAACGGCATGG 5363
 Db 5353 GGAAACAGCTACTACAGCTCTGTAAATATGTTACCTGAGGCTGATGTATGTTAATGGATGG 5412
 Qy 5364 AGTGGGCTCTGAGACTGAGCCCTGCTGTGGCTGGCACCGTCAACCCACCTGGGCA 5423
 Db 5413 GTATCAGCTTCACAGCAGCCCATGTCCTAGCGGGCACCATCACCCCACTTGGAC 5472
 Qy 5424 AGAGAAATGTCAAGCTGCCATGACACAGCCCTCAACCTGTGTGAGTGGCGCCAGGGCA 5483
 Db 5473 GCTGCAACATCTCCCTGCTATGAGAAATGCTTAAACTCCATTTAGTGGCGCTAAGAA 5532
 Qy 5484 AAGACAGCTCGGGCCAGGTCACTGTCTTTGGGCGCCGCTGCGGTGACAAACGAA 5543
 Db 5523 AGGAACAGATTAAAGGCAAGTCAACATCTTTGGCAGGAAGCTCCGGTCCATGGAAGAA 5592
 Qy 5544 ATCTCTATCTCTGAGCTTTGATCGCGTAAACACGACAGAGAAATCTATGATGACCA 5603
 Db 5593 ATCTCTTCTCATTTGACTATGATCGAATATTCGACTGAAAGATCTATGATGACCA 5652
 Qy 5604 GCAATTCACCTTCGGATCTGTAGACACAGCGGGCGCCAGGCTCTGTGTCACCCA 5663
 Db 5653 GGAATTCACCTTCAGGATCATTTATGACAGAGTGGCGCGCCCTTCTCTGGTGGCCA 5712
 Qy 5664 GCAGCAGCTGAATGGTGTCAACGTGACATCTCCCTGGGGTTTACATGCTGGCATCC 5723
 Db 5713 GCAGCGGCTGGAGCTGTCAACGTGTCTACTTCTCAATGGGCGCTGGCTGGGCTTC 5772
 Qy 5724 AAGGGGATCATGTCTGAAAGATGGAATACGACAGGGCGGCGCATCACATCCAGGA 5783
 Db 5773 AGCGTGGGCTCATGAGGAGAGGACAGACATCGAACAGAGGCGCATCGTGTCCCGCA 5832
 Qy 5784 TCTTCGCTGATGGAGACATGAGCTACACATCTTAGAGAGTCCATGGTGTGTGCTAC 5843
 Db 5833 TGTTCGCTGAGCGGAAAGTGGAGTACTCTCTACCTTGACAGTCCATGGTCTCTCTGC 5892
 Qy 5844 TACACAGCAGAGGCGAGTATCTTTGAGTTGCAAGAAATGACCGGCTCTCTTCTGTGA 5903
 Db 5893 TTCAGAGCAACGTCAGTATATTTTGGATGATGACTCTCTGACCGCTCTCTGCGCTCA 5952
 Qy 5904 CGATGCCAAAGTGGGCGGAGACACTAGAGACCATCCGCTCAGTGGGCTTACTACAGAA 5963
 Db 5953 CCATGCCAGGCTGGCGGACAGCATGTCCACACACACCTCCATCGGCTTACTACCGTA 6012
 Qy 5964 ACATCTATCAGCCCTGAGGGCAATGCTCAGTCATACAGGACTTCACCTGAGGATGGC 6023
 Db 6013 ATATTACAAACCGGCTGAAAGCAATGCTTCGTCATCTTTGACTACAGTATGACGGCC 6072
 Qy 6024 ACCTCTTTCACACCTTCTACCTGGGCACTGGCCGAGGGTATATACAAGTATGGCAAA 6083
 Db 6073 GCATCTCAAGACCTCTCTTTTGGGCAACCGGACCGAGGTTCTACAAGTATGGGAAAC 6132
 Qy 6084 TGTCAAACTGGCAGAGAGCTCTATGACACCAACAGGTCAGTTTCACTATGACGAGA 6143
 Db 6133 TCTCAAGTTATGAGAGTTGCTTACGACAGTACCGCGCTCACCTTCGGGTATGACGAGA 6192
 Qy 6144 CGGAGGATCTGGAAGACCATCAACCTTACAGAAATGAGGGCTTCACCTGACCACTCCGCT 6203
 Db 6193 CCATCTGGTCTTGAAGATGTCACCTCCAAAGTGGGGCTTCTCTGCAACCATCAGT 6252
 Qy 6204 ACCGTGAGTTGGGCGCTGATGACCGACAGATCTTCGCTTACCTTACCTGAGAGGCTAG 6263
 Db 6253 ACCGGAAGTTGGGCGCTGTTGGGACAGGAGATCTACAGGTTCTCGAGAGGAGCTAG 6312
 Qy 6264 TCAACGCGCTTTTGACTACAACT---ATGACAAACAGCTTCCGGGTGACAGCATGAGG 6320
 Db 6313 TCAATGCCAGTTTGTACTACACTTATCATGACACAGCTTCCGATCGCAAGCATCAAGC 6372
 Qy 6321 CTGTGATCAACAGAGACCCCACTGCCATTTGATCTCTATCGTATGATGATGTGTGAGCA 6380
 Db 6373 CCGTCATAAGTGAGACTCCCTCCCGTTCACCTTACCGCTATGATGATGATGATTTCTGCA 6432

QY 6381 AGACAGCAGTTTGGGAAGTTTGGTGTCTTACTATGACATTAACCAAGATCATCACCA 6440
DB 6433 AGTGGAAACATTTTGGTAAAGTTTGGAGTCACTTATTAAGACATCAACCAAGATCATCACCA 6492
QY 6441 CAGCTGTATGACCCACCAACCAAGCATTTTGTATGATGCAATGAGGATGAAGGAAGTGCAGT 6500
DB 6493 CTGCGGTATGACCTCAGCAACACATTCAGACACCCATGGCGGATCAAGGAGTCCAGT 6552
QY 6501 ATGAGATCTCCGCTCGCTCATGTATGATGACCGTCCAGTATGATTAATCATATGGGGGAG 6560
DB 6553 ATGAGATGTTCCGGTCCCTCATGTATGATGACCGTCCAGTATGATGACGATGGCGAGG 6612
QY 6561 TAGTGAAGAAGAGCTGAAGTATGAGCCCTAGCGCAATACACATCGGTCTCTCTATGAGT 6620
DB 6613 TGATCAAGAGGAGCTAAACCTGGGGCCCTATGCAATACCAAGAGTACACCTATGACT 6672
QY 6621 ATGATGTGACGGCAGCTGACAGAGCTTCCATCAATGACAAAGCCACTCTGGCGGTACA 6680
DB 6673 ACATGGGACGGGAGCTCCAGAGCGTGGCGCTCAATGACCGCCGACCTGGCGGTACA 6732
QY 6681 GCTACGACCTCAATGGGAACTGACATCTAGCCCTGGGAGAGTGCAGGCTACAC 6740
DB 6733 GCTATGACCTTAAATGGGAATCTCACCTTACTGAACCCAGGCAACAGTGTGGCGCTATGC 6792
QY 6741 CACTACGGTATGACATCCGACCGCATCACTCGGCTGGGTGAGTGCATCAAGATGG 6800
DB 6793 CTTGGCGTATGACCTCCGGATCGGATACCAAGACTCGGGATGTGCAGTACAAATTTG 6852
QY 6801 ATGAGGATGGCTTCTGAGGAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGCC 6860
DB 6853 ACAGCATGGGTATCTGTGCGAGAGGGGTCTGACATCTTCGAATACAAATTCGAAGGGCC 6912
QY 6861 TGCTCATCAAGCGCTTACAAACCGGCTGGCAGCTGGAGTGCAGTACCGTACGATGGCC 6920
DB 6913 TCCTAAAGAGCTTACAAAGGCGAGCGGTGAGTGTCCAGTACCGTATGATGGCG 6972
QY 6921 TGGGCGCGCGTGTCCAGCAAGAGCAGCCACAGCCACCGTGCAGTCTTCTATGACAG 6980
DB 6973 TAGGACGGCGGCTTCTTACAAAGCAACCTGGGCGCACCACTGCAGTACTTCTACTCTG 7032
QY 6981 ACCTGACCAACCCACCAAGCTCACCAACCTGTACACCACTCCAGCTCTGAGATCACCT 7040
DB 7033 ACCTCCCAACCCGAGCGGATACCCATGTCTCAATCTCTCAATCTCGGAGATTAACCT 7092
QY 7041 CCCTCTACTAGCACTTGAAGGACACCTCTTTTGGCATGGAGCTGAGCAGTGGTATGAGT 7100
DB 7093 CACTGTACTAGCACTCCAGGCGCACCTCTTTGCCATGGAGCAGCAGTGGGAGGAGT 7152
QY 7101 TTTACATAGCTTGTGACAAATCGGAGCCCTCTGTGTCTTTAGTGGACAGGTTTGA 7160
DB 7153 ACTATGTTGCTCTGATAACACAGGGACTCTCTGGCTGTGTTTCCAGCATCAACCGGCTCA 7212
QY 7161 TGATCAAGCAAACTCTGTACACAGCCTATGGGAGATCTACATGGATACCAACCCCAACT 7220
DB 7213 TGATCAAAACAGCTGCAGTACACGGCTATGGGAGATTTATATGACTCCAAACCCGACT 7272
QY 7221 TTCAGATCATATAGGCTACCATGGTGGCCTCTATGATFCCACTCAACAGCTTGTCCACA 7280
DB 7273 TCCAGATGCTATGGCTTCCATGGGGACTCTATGACCCCTGACCAAGCTGTGCTCACT 7332
QY 7281 TGGGCGCGGAGTTATGATGTGCTGGCGGAGCGTGGACTAGCCCAAGCAGCAGTGT 7340
DB 7333 TCACCTACGCGTATATGATGTGCTGGCAGGACGATGGACCTCCCGAGCTATACCATGT 7392
QY 7341 GGAAGCAGCTTAGTAGCAGCAACGTCTATGCTCTTAAATCTCTATATGTTCAAAAACAAACA 7400
DB 7393 GGAAGAACGT---GGGCAAGGAGCGGCGCCCTTTAACTGTATATGTTCAAGAGCAACA 7449
QY 7401 ACCCATCAGCACTCCAGGACATCAAGTGTCTCATGACAGATGTTAAGCTGGTGC 7460
DB 7450 ATCCCTCTCAGCAGTGAAGTATTTGAAGAACTAGTGACAGATGTGAAGAGCTGGCTTG 7509
QY 7461 TCACCTTTGGATTCCAGCTACACACAGTGAATCCCTGTTTATCCCAAAACAGACATGGATG 7520

DB 7510 TGATGTTGGATTTAGCTTAGCAACATCATCTCTGGCTTCCGAGAGCCAAATGTATT 7569
QY 7521 CCATGGAACCCCTCTACAGAGCTCATCCACACAGATGAAGAAACGAGGAGTGGGCAACA 7580
DB 7570 TCGTGCCTCTCTCTATGA-----ATTGTCAAGAGTCAAGCAAGTGAAGATG 7617
QY 7581 GCAAGTCTATCTCGGGTACAGTGTGAAGTACAGACAGCTCAAGGCTTTGTCACT 7640
DB 7618 GACAGCTATTACAGGTGTCACAGAAACAGAGAGACATTAACAGGCTTCATGGCTC 7677
QY 7641 TAGAAGCGTTTACCAGCTCTATGGCTCCAAATCACCAGCTGCCAGGAGTCCAAAGA 7700
DB 7678 TGGAA-----GGACAGGTCAATTAATAAAGCTCCAGCCAGCATCCGAGAGAAAGCAG 7731
QY 7701 CCAAGAGTTTTCATCTCAGCGCTCAGTCTTTGGCAAGGGGTCAAGTTTGGCTTTGAAG 7760
DB 7732 GTCATGTTTCCCAACACCGCTACTACCGCCATCATTTGGCAAGGCTATGTTTGCATCAAG 7791
QY 7761 ATGGCGAGTGAACCAAGACATCATCAGTGTGCCAATGAGATGGGGAAGGGTTGCTG 7820
DB 7792 AAGGGCGGTGACCAAGCGGTGTCCAGCATCGCCAGCAAGATAGCCGCAAGTGGCAT 7851
QY 7821 CCATCTTGAACCATGCGCCACTACTAGAGAACCTGCACTTACCAATTGATGGGGTGATA 7880
DB 7852 CTGTGCTGAACACCGCTACTACTGGCAAGATGCACTACAGCATCGAGGGCAAGGACA 7911
QY 7881 CCATTTACTTTGTAAACCCAGACCTTCAAGAGGTGACCTGGCCATCTCTGGGCTCAGTG 7940
DB 7912 CCCTACTTTTGTGAAGATTGGCTCAGCGATGGGCACTGTGTCACTAGGACCAACA 7971
QY 7941 GGGGCGCGCAACCTTGGAGATGGGTCAACGTCACTGTGCCAGATCAACACAGTAC 8000
DB 7972 TCGGCGCAAGTGTCTAGAGAGCGGGTGAACGTGACCGTGTCCAGCCACGCTGCTGG 8031
QY 8001 TTAATGGAGGAGTACAGCTACAGAGATCCAGCTCCAGTACGGGCACTGTGCTTGA 8060
DB 8032 TCAACGGCAGGACTCGAAGTTTCAAGAACTTCAAGTTCAGTACTCCAGTACTCCAGCTGCTCA 8091
QY 8061 ACACAGCTTACGG-----GACAACTTGGATGAGGAGAGGCAAGGCTCTGGAGC 8111
DB 8092 GCATCGCTATGGCTCTACCCCGCACCTTGAAGAGAGAGGCGCCGCTCTGGACC 8151
QY 8112 TGGCCCGGAGAGAGCGGTGGCGCAAGCGTGGCGCGCGAGCAGAGACACTTGGCGGAG 8171
DB 8152 AGCGGAGACAGAGGGCGCTGGGCGCGGCTGGGCAAGGAGCAGCAAGGCAAGGCGAG 8211
QY 8172 GGGAGGAAGCGCTGGGGGCTTGACAGAGGGGAGAGCAGCAGGCTGTGAGCACAGGCG 8231
DB 8212 GGAGAGGGGAGCGCTGTGAGCTGAGGGGAGAGCAGCAGCTTCTGAGCACCGGCG 8271
QY 8232 GGTGCAAGGCTACGAGCGCTTTTTCGTGATCTTGTGAGAGTACCAGAACTGTCTAG 8291
DB 8272 GCGTGAAGGGTACGAGGGATATTACGTCTTCCGTGGAGCAATACCCAGAGCTTGCAG 8331
QY 8292 ACAGCCCAACACATCCACTTATGAGAGAGGAGATGGCGGAGGTGACAGA 8348
DB 8332 ACAGTAGCAACATCCAGTTTAAAGAGAAATGAGATGGGAAGAGGTAAACAAA 8388

RESULT 12
AAK51828
ID AAK51828 standard; cdna; 13202 BP.
XX
AC AAK51828;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 373.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW	nervous system disorder; arthritis; inflammation; ss.	QY	1177	GATCAGGAGGACAGCCAGCAGTGGCTGTCGCAACCGACGCTCTCCCTATATCCCTC	1236
XX		DB	1464	AAATGGAGTTAGCAAGGAGACAGGGGACCGAGTCCATGACACTACTTACTCTCCAAT	1523
OS	Homo sapiens.	QY	1237	AGGGGCACTGGCTTAGAGACCCCTGACAGGAAGGCAAGAAACCAAGAGAAAGCC	1296
PN	WO200157190-A2.	DB	1524	TGGAGGAAAGTTTCTGTATAAATCAGAGAAAGAAAG-----	1558
XX	09-AUG-2001.	QY	1297	CAGTAGTTTCTTTCCAGAGGACAGTTTTCATAGATTCTCGGAAATTTGATGTGGGAGCG	1356
PF	05-FEB-2001; 2001WO-US004098.	DB	1559	-----TGTTTCAGAGGGACCGGCGATAGACACTGGAGAGTTGACATTTGGTGACA	1610
XX	03-FEB-2000; 2000US-00496914.	QY	1357	AGCCTCCAGAGAGTTCTCTCGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCC	1416
PR	27-APR-2000; 2000US-00560875.	DB	1611	GGTCATGCAGACCAATCCACCTGGTTTATCTGGCGTTTCAGATTACTATCCACCATCC	1670
PR	20-JUN-2000; 2000US-00598075.	QY	1417	TGTGCATCTGAATTCATATGTCTCTCGGAAGGACGACCTGTTGGCATTTATGCGAG	1476
PR	19-JUL-2000; 2000US-00620325.	DB	1671	AAATATCTGAAGTTCAATATTTCTTTAGCCCAAGGACTCTCTGCTGGGAATTTATGCGAG	1730
PR	01-SEP-2000; 2000US-00654936.	QY	1477	AAAAGGCTCCCTCTCCATACATACAGTTTTCAGATTCTGAGCTGCTGGATGCGAGAG	1536
PR	15-SEP-2000; 2000US-00663561.	DB	1731	AGAAACATTCACCTACATACACTCTCAGTTTGTAACTAATGATGSCAAACA	1790
PR	20-OCT-2000; 2000US-00693325.	QY	1537	GCTCCTAACCCAGGAGGCGGAGCTTAGAGGGACCCCGCGGAGTCTCGGGGAATGT	1596
PR	30-NOV-2000; 2000US-00728422.	DB	1791	GCTGCTCAAGCAGGACTCCAAAGGGCTCTGATGATACACAGCACTCCCTCGGAACTGAT	1850
XX	(HYSE-) HYSEQ INC.	QY	1597	GCCCCCTCCAGCATGAGACAGGCTTCATCCAGTATTTGATTCAGGAATCTGSCACTT	1656
XX	Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y;	DB	1851	CTTAATCTCGCTTCAGGAGCAGGTTTCTAGAGATATATGATCAAGACCTTGATATCT	1910
PI	Ma Y, Zhao QH, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;	QY	1657	GGCTTTTACAAATGACGGAAGGAGTCTAGAAAGTGGTTTCTTTCTACCACTGCCATTGA	1716
PI	Xue AJ, Yang Y, Wejhrman T, Goodrich R;	DB	1911	GGCGTTTACAAATGATGAAAGGAGTGGAGCAAGTATTCGTGTTAACTACAGCAATGA	1970
XX	WPI; 2001-476283/51.	QY	1717	GTCGGTGAATACCTGCCCGCACTGCTATGGCAATGGTGTACTGCACTCTGAGACTG	1776
DR	P-P5DB; AAM78695.	DB	1971	AATAATGATGACTGTTCAACCAATTCGAATGGAATGGAGAGTGTATCTTGCCCATG	2030
PT	Nucleic acids encoding polypeptides with cytokine-like activities, useful	QY	1777	CACTGCTCTCTGGGTTTCTCGGCGCCGACTGTGGCAGAGCCCTCTGCCCCGCTCTG	1836
PT	in diagnosis and gene therapy.	DB	2031	TCATTGTTTCCAGGATTCCTTGACCTGACTGTGCTAGAGATTCCTGCCCTGTGCTGTG	2090
XX	Claim 1; Page 1414-1426; 6221pp; English.	QY	1837	TAGCGAAATGCCAATACATGAAGGAGATGCTTGTGCCACAGTGTGCTGGAAGGCGC	1896
CC	The invention relates to polynucleotides (AAK51456-AAK53435) and the	DB	2091	TGGTGGAAATGGAGAAATACGAGAAAGGACACTGTCTGTCCGGGATGCTGGAAGGGCC	2150
CC	encoded polypeptides (AAM78323-AAM80302) that exhibit activity elating to	QY	1897	TGAGTGCCATGTGCCCAACCACTGATGATGTGCGCTGCGCTGACAGCAACATGACACTG	1956
CC	cytokine, cell proliferation or cell differentiation or which may induce	DB	2151	AGAGTGTGACGTTCCGGNAGAACAAATGATGATGATGATGATGATGATGATGATGATG	2210
CC	production of other cytokines in other cell populations. The	QY	1957	CATCAGGCGCACTGCTGCAACCTTGTGCTACAGGGGAGAGCTGTGAGGAAGTGA	2016
CC	polynucleotides and polypeptides are useful in gene therapy, vaccines or	DB	2211	CATCATGCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2270
CC	peptide therapy. The polypeptides have various cytokine-like activities,	QY	2017	CTGCATGCAACCCCATGTTTCAGGCGGGGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2076
CC	e.g. stem cell growth factor activity, haematopoiesis regulating	DB	2271	CTGCCTAGACCCCAATGTTTCCAAACCATGGCATCTGTGTAAGAGGAATGTCTACTGTTT	2330
CC	activity, tissue growth factor activity, immunomodulatory activity and	QY	2077	TGTGGGATGGGAGGACCAACTCGAGACCCCGAGGCGCCACATGCTTAGACAGTGTTC	2136
CC	activation/inhibin activity and may be useful in the diagnosis and/or	DB	2331	TACTGGCTGGGAGGAGTTAACTCTGAAACACCACTTCTCTGTATGTCAAGAGCAGTCTC	2390
CC	treatment of cancer, leukaemia, nervous system disorders, arthritis and	QY	2137	AGGCCAGGAACCTTCTCCCGGACACCGGGCTTTGCACTGTGTACCAAGCTGGAAGTGA	2196
CC	inflammation. Note: Records for SEQ ID NO.2110 (AAK52581), 2111	DB	2391	AGGACACGGAATCTTCTTCTGAGCGTGTGAGTATGAGTGTGATCCCAAGTGGACAGG	2450
CC	(AAK52582) and 3666 (AAM8020) are omitted as the relevant pages from the	QY	2197	ACAGACTGTTTCTATCGAGATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2256
CC	sequence listing were missing at the time of publication	DB	2451	ATCTGACTGCTCAACAGAGCTGTGTACCATGAGTGTGTGTAGCCATGAGTGTGTCTCAAG	2510
XX	Sequence 13202 BP; 3921 A; 2764 C; 2865 G; 3652 T; 0 U; 0 Other;	QY	2257	GGGCACCTGCGCTGCGAGGATGCTGATGGGGGACGCTGCGACAGCGGGCTGCCA	2316
SQ	Query Match 39,28; Score 3274.2; DB 4; Length 13202;				
	Best Local Similarity 65,78; Pred. No. 0;				
	Matches 4910; Conservative 0; Mismatches 2498; Indels 67; Gaps 7;				
QY	880 GGGCACTTCTCTTCAAGCTGGA---GGCACTCCCGCTCTTCTGCAACATCACC	936			
DB	1164 CAGGCAATTCCTGTTCAACATGATGCTGTTCTCTGCGATCTTCAGTGCAGCCAGTCA	1223			
QY	937 AGGCTACCCACTGACGTCAGCAGTACTCTCTCCGCGCGGACCTTCCCGCGAG	996			
DB	1224 GAATACCTCTGACATCCAAATACCGGTACTCGCCCTTCCAGGCTTCTCTCGAAG	1283			
QY	997 CACCTTCGCGCGCGCGCTTTAACTTCAAGAGCCCTCCCAAGTACTGTAACCTGGAAGTG	1056			
DB	1284 CACCTTTCCGACCTCCCTTTACCTTTAAACAACTTACAGGTGCTGCAACTGGAAGTG	1343			
QY	1057 CGGAGCCCTGAGCGCATCTGCTATCTGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTG	1116			
DB	1344 CACAGCAATGAGCGCCACTGCAATACAGTGTATTTGGCTTGTACTAGCCCTATGTGAT	1403			
QY	1117 GGCCATGCACTGTTTGGCTTAACTGGCACTGCGACCGCATGGAGGGGAGATGATGA	1176			
DB	1404 TGCAGTGCATTTGTTGGCTGACTTGGCAGTTGCAACAGTTGAAGGAGAGCTGTATGC	1463			

Db 2511 AGGAATTTGCCAGTGTGAAGAAGCTGGTAGGACCAACATGTGAGAAAGCTCCTGTCA 2570
 Qy 2317 CCCGGCTGTGCCAGCATGGAGCTCTGCGGAGCGCAAGTGCAGTGCAGCGCTGTGCTG 2376
 Db 2571 TTCTCATTTGACTGAGCATGGCCAAATGCAAGATGGAATGTGAGTGTAGCCCTGGATG 2630
 Qy 2377 GAATGGGGAACACTGCCACCTGCTCTACTATCTGGATAGGGTATGTTAAAGAGGGTTGCC 2436
 Db 2631 GGAGGGGACCACTGCACAAATGTCTACTACTTATAGT---GCTGTCCGAGATGGCTGCC 2687
 Qy 2437 TGGGTTGTGCAATGGCAACGACATGATCCTTAGACCTGAATGGTTGGCACTGCCCTG 2496
 Db 2688 AGGGCTCTGCTTTGGAAATGACATGTACCCCTGGATCAAAATGTTGGCACTGTGTGTG 2747
 Qy 2497 CCAGCTGGGCTGGAGAGAGCTGCTGTGACACTTCCATGGAGACTGCCCTGCCGTGACAG 2556
 Db 2748 TCAGGTGGTTGGAGTGGGACAGCTGCAATGTTGTGATGAAATGCTTTGTGGAGATAA 2807
 Qy 2557 CAAGAACAATGATGGAGATGCCCTGTGGTGAATGACATGGACCCCTGACCTGCTGCCAGCC 2616
 Db 2808 CTTGGACAATGATGGAGATGTTTAAACGACTGTGTGGATCCTGACTGTTGTCAACAAAG 2867
 Qy 2617 CCTGTGCCATATCAACCGCTGTGCTTGGCTCCCTTAAACCTCTGGACATCATCCAGGA 2676
 Db 2868 CAACGTGTTATTAAGTCTCTCTGCCAGGCTCACAGATCCTCTTGAACCTCATTCASCA 2927
 Qy 2677 GACACAGTCCCTGTGTCACAGCAGAACCTTACATCCCTTATGACCCGATCAAGTTCCCT 2736
 Db 2928 AAGCCAAATCTCTTCTCAGCACACTTCAAGACTTTTATGATCGAATCAAAATTCCT 2987
 Qy 2737 CGTGGGAGGACAGCGACATATCCCGGGAGAACCCCTTTGATGGAGGGCATGC 2796
 Db 2988 CATTGGCAAGACAGTACTCATGTCAATCTCTGAGGTGTCATTTGACAGGCGCTGC 3047
 Qy 2797 TTGTGTTATTCGTGCCCAAGTGAATCAGATCAGATGGAAACCCCTCGTGTGGTGTGAACAT 2856
 Db 3048 CTGTGTGATTGAGGCCCAAGTGTGGCCATAGATGGAACTCCTCTAGTGGGAGTGAATG 3107
 Qy 2857 CAGTTTGTCAATTAACCTCTCTTTGGATATACATCAGCAGGCAAGATGCGACTTTGA 2916
 Db 3108 CAGTTTCTTGACCAACAGTGAATGAGGTTTACCATCAGCGGCAAGATGGAAGCTTTGA 3167
 Qy 2917 CTTGTGTAAATGGCGCATCTCCATCATCTGCGGTTCAGCGGGCAGCTTTTCATCAC 2976
 Db 3168 CTTGTGCGCATCGTGGCATCTCTGTCACTCTTAATCTTGACCGATCCCTTTCTCGCC 3227
 Qy 2977 ACAGGACACACCTGTGGCTGCCATGGGATGGCTCTTTGTCTATGGAAACCATCATCAT 3036
 Db 3228 TGAGAAGAGAAACACTCTGTTGCTTGGAAATCAGTTTATTTGTGTAGAGAAAGTCAACAT 3287
 Qy 3037 GAGACATCAGGAGATGAGATTCAGCTGTGACCTGAGCAATTTTGGCCGCCCAACCC 3096
 Db 3288 GCAGAGATTGTATCAGACCCGCCATCTCGGATATCTCCAACTTTATCAGCCCAACCC 3347
 Qy 3097 AGTCGTCTCTCCATCCCACTGACGTCTCTGCCAGCTCTCTGTGCAGAGAAAGCCCAT 3156
 Db 3348 TATTGTGCTTCTTCAACCGCTCACATCATTTGGAGGGTCTGTCCAGAGAGGGAACTAT 3407
 Qy 3157 TGTGCCGGAATTCAGGCTTTGAGGAGGAAATCTCTATCTCTGCTGTCAGATGAGGCT 3216
 Db 3408 TGTTCCTGAGCTGACAGTTGTACAGGAGGAAATTCCTTCTCCCTCCAGCTTTGTGAGGCT 3467
 Qy 3217 GAGCTACCTGTAGCAGCCGACCCCTGGCTACAAATCTGTCTGAGGATCAGCCCTCAGCCA 3276
 Db 3468 GAGTTACCTGAGCAGCCGACCCCTGGGTATTAACCCCTGCTACGGATCCTTCTGACACA 3527
 Qy 3277 CCCGACCATTCCTTTCAACCTCATGAAGGTGACCTCATGTGTAGCGGTGGAGGGCGGCT 3336
 Db 3528 TTCAACGATTCCTGAGGATGATATAAGTAGTACACCTCACAGTAGCTGTGGAAGGCGACT 3587
 Qy 3337 CTTCAGGAGTGTGCTGTCAGCCCGACAGCTGCTCTATTATTCTTTGGGCAAGAC 3396

Db 3588 CACACAGAGTGGTTTCCCGCGCAATTAATCTTGTCTACACATTTGCTTGGAAACAAGAC 3647
 Qy 3397 AGACGTCTTACAAACCAAGAGTGTTTGGGCTTTTCAAGCCTTTTGGTTCOGTGGGTATGA 3456
 Db 3648 CGATATCTATGACAGAGAGTGTGGGCTGTGCGAGAGGCTTTGGTATCTGTGGGATAGA 3707
 Qy 3457 ATATGAATCTCTCCAGATCTAATCTGTGGGAAAAAAGAACAAACAGTCTGCAGGGCTA 3516
 Db 3708 ATATGAACCGTCCCTGACTTTATTTCTCTGGAGCAAGAGGACAGTCTGTTTCAAGGTT 3767
 Qy 3517 TCAAAATGACGGCTCCAAAGCTTTGGAGGATGGAGCCTAGACAAACATCATGCCCTCAACAT 3576
 Db 3768 TGGATGATGCTTTCTAACCITAGGAGCTGGTCTTTTGAATAGCATCACATTTGAATCC 3827
 Qy 3577 TCAAGTGTGTATCTGCAAAAGGAAATGGGGAGAACAGTTTGTCTCAGCAGCCTCC 3636
 Db 3828 TCAAGTGAATATACATAAAGGAAATGGAGAAATATGTTTCATTTCCAGCAGCCGCC 3887
 Qy 3637 TGTCTATGGAGCATCATGGCAATGGGCGCGGAGAGACATCTCTGCCCCAGCTGCAA 3696
 Db 3888 AGTCATATCAACATAATGGGTAAATGGACACCAAGAGAGTGTAGCTTGACCACTGCDA 3947
 Qy 3697 CGGCTTGTGACGGCAACAAAGCTCTGCGCCCAAGTGGCCCTCACTGTGGTCTGACGG 3756
 Db 3948 TGGCCAGCCCAACAACAACACTCTTGTCTCTGTCGCTTAGCTTCTGGCCCTGATGG 4007
 Qy 3757 GAGCTCTATGTGGTGAATTTCAACTATTAAGAGGATCTTCCCTCTGGAATATGCAC 3816
 Db 4008 CAGTGTGTATTTGGGACTTCAATTTTGAAGGAAATATTTTCCCTCGGAAACTCCCT 4067
 Qy 3817 CAACATCTTAGAGCTGAGGAATAAAGATTTTACAGATAGTACAGTCCAGCAGCAAAATA 3876
 Db 4068 TAGTATTTTGAATTAAGCACA-----AGTCTCTCACAAATA 4106
 Qy 3877 CTACCTGGCCACAGACCCCATGAGTGGGCGCTCTTCTTCTGACAGCAACAGCGGCG 3936
 Db 4107 CTATCTGGTATGGACCTCTGTCTGAATCACTCTATCTATCAGACACCAATCTCGCAA 4166
 Qy 3937 GGTCTTTAAATCAAGTCCACTGTGTGTGAGAGACCTTGTCAAGAACTCTGAGGTGT 3996
 Db 4167 AGTCTACAGATTGAATCTTGTGTGAGACGAAGATCTGTCCAGAAATTTTGAAGTGT 4226
 Qy 3997 TGGGGGACAGTGTACAGTGTCTCCCTTTTGAATGACACTCGCTGGGGGATGTGGGAA 4056
 Db 4227 GGCAGGAATCTGTGTGATGAGTGTCTTCCCTTTGACAGAGTCTATGTGAGATGGTGGAG 4286
 Qy 4057 GGCACAGAGCCACACTCACCAATCCAGGGGTATTACAGTGGACAAAGTTTGGGCTGAT 4116
 Db 4287 AGCATCGGAGCTTCTACTGAATAGCCCTCAGGCAATCAGATTGATAGCATGGATTTAT 4346
 Qy 4117 CTACTTCTGGATGGCACCATGATCAGACGCAATCGATCAGAAATGGGATCATCTCCACCT 4176
 Db 4347 TTACTTTTGGATGGGACTATGATTCGCAAAATTTGATGAGAATGCTGTGATCACAATGT 4406
 Qy 4177 GCTCGGCTCTAATGATCTCACATCAGCCCGGCACTCAGCTGTGATCTGTATGATAT 4236
 Db 4407 AATCGGCTCAATGGTCTGACTTCCACCAACACTGAGCTGTGATCAGGAATGACAT 4466
 Qy 4237 TTCCCAAGTAAAGACTGGAGTGGCCCAAGACTTAGCCATCAACCAATGGACAACTCACT 4296
 Db 4467 CACTCAGTGGCATTAGAGTGGCCACAGACCTTTCAGTAAATCTTATGACAAATTCAT 4526
 Qy 4297 TTAGTCTCTGACAAACAATGTGGTCTGCAATCTCTGAAATCTCTGAAACCAACAGGTGGCAT 4356
 Db 4527 GTATGCTTGGATAACAACATTTGTGTGCAAAATTTCTGAGAACAGGCGTGTTCGGATCAT 4586
 Qy 4357 CGCGGGAGGCCCATCTCAGCTGCCAGTCTCTGCAATTTGACCTTCTGTAAGCAAGGT 4416
 Db 4587 CGCAGGACGCCCATTTCACTGCCAGTGTCCAGGCAATCGATCATTTCTCTGTGACAGAGT 4546
 Qy 4417 GGCCATCCAGCAACCTGTGAGTCAAGCCCGCTTTGGTGTGTTTCAACAATGGGTCCT 4476
 Db 4647 ASCAAATTCACCTCTAGAGTCAGGAGGCGCATCAGTGTCTCCACAGCGGCTGCT 4706

QY	4477	GTATATTGCTGAGACTGATGAGAAAAAGATCAACGGGATCAGGCAGGTCAACCACTAGTGG	4536
DB	4707	CTTCATAGCTGAAACAGACGAGAGGAAAGTAAACCGCATTCAGCGAAAGTACCCACCAATGG	4766
QY	4537	AGAGATCTCACTCGTTGCTGGGGCCCCAGTGGCTGTGACTGTAAAAATGATGCAACTG	4596
DB	4767	GGAGATCTACATCATCGCTGTCGCCCACTGACTGTGACTGCAAAATTGATCCAACTG	4826
QY	4597	TGATTTGTTTTCTGGAGACGATGGTTATGCCAAGGATGCAAGTTAAATACCCTATCTTC	4656
DB	4827	TGACTGTTTTTCAGGTGATGTGCTGCTATGCCAAAGATGCAAAAGATGAAGACCCCTTCCTC	4886
QY	4657	CTTGGCTGTGTGCTGATGGGGAGCTCTACGTGCGCCGACCTTGGGAAACATCGGAATTCG	4716
DB	4887	CTTAGCAGTGTGCGCTGATGGAAACCCCTCTATGTGGCAGACCTCGGAAATGTTGGAATTCG	4946
QY	4717	GTTTATCCGGAAAGAACAGCCCTTCTCAACACCCAGAAATGATPATGAGCTGTCTCAACC	4776
DB	4947	TACCATCAGCAGGAACCAAGCCCACTGAATGACATGAAGATTATGAGATTTGTTCAACC	5006
QY	4777	AATTGACCAGAGCTCTATCTCTGTTTGATACACCGGCAAGCACCTGTACACCCCAAGCCT	4836
DB	5007	CGCTGATCAGAACTGTACCAAGTTCACTGTAAATGGAAACCCACTACACACCTGAACTT	5066
QY	4837	GCCCAAGGAGACTACCTGTGTAACAATTCACCTACACTGGGAGCGGCACATCACTCAT	4896
DB	5067	GATAACAAGGACTATGTTTATACTTCACTACAAATCTGAAAGTGACTTGGGCGCGAT	5126
QY	4897	CACAGACAACTGGCACATGTTAAATGTCCGCGAGACTCTACTGGGATGCCCTCTG	4956
DB	5127	TACCAGCAGCAATGGCAATTCAGTGCAATTCGGCCGTGATCGAGCGGGAATCGCGTATG	5186
QY	4957	GCTGGTGTCCAGATGCCAGTGTACTGGGTGACCATGGGACCAACAAAGTGCATCTCAA	5016
DB	5187	GCTTGTGGTGCTCGCGGACAAGTATACTGGCTGACTATAAGCAGCAATGGAGTCTGTAA	5246
QY	5017	GAGTGTGACCAACAAGGACAGAGTTGGCCATGATGACATACCATGGCAATTCGGGCT	5076
DB	5247	AAGAGTGTGAGCCCAAGGCTATAATCCGGCTTAAAGACCTATCCAGGAAACACAGGGCT	5306
QY	5077	TCTGGCAACCAAAAGCAATGAAAACGGATGGACAAACATTTTATGAGTACGACAGCTTTGG	5136
DB	5307	TCTGGCTACCAAAAGTAAACGAAATGATGGACACACCGTTTATGAGTATGACCCCGAGGG	5366
QY	5137	CCGCTGACAAATGTGACCTTCCCTACTGGCAGGTGAGCAGTTTCGAAAGTGATACAGA	5196
DB	5367	ACACCTGACCAATGCAACGTTTCCACTGGAGAGGTGAGCAGCTTCCACAGTGCACCTGA	5426
QY	5197	CAGTTCAGTGATGCTCCAGGTAGAGACCTCCAGCAAGGATGATGTACCATAACCAACCAA	5256
DB	5427	GAACCTGACAAAAGTGGAGCTAGATATCTCCAACCGTGAAAATGTCTCATGTCAACCAA	5486
QY	5257	CCTGTCTGCTCAGCGGCTTTCTACACACTGCTGCAAGACCAAGTCGGGAACAGTACTTA	5316
DB	5487	CTTGACGGCAACTAGTACCATATATATTTTAAAAACAGAAAAATACTCAAAGTACCTTACG	5546
QY	5317	CATCGGGCCGATGGCTCTTGGGGCTGCTGTGGCCACGCGATGGAGTGGCGCTGCA	5376
DB	5547	GGTGAATCCAGATGGTTCCCTGGGTGTCATCTTTGACGCGGATGGAGATCGGCTTCAG	5606
QY	5377	GACTGAGCCCACTTGTGTGCTGGCACCGGTCAACCCCAACCGTGGGCAAGAGGAATGTCA	5436
DB	5607	CTCAGAGCCCAACATCTGTGGCAGGGGAGTCAACCTCATCCCTGGGCAAAATGCAACATCTC	5666
QY	5437	GCTGCCCATCGAACCGGCTCAACCTGTGTGGAGTGGCGCCAGCGCAAGAGCAGGCTCG	5496
DB	5667	ATTGCCCGGAGACCAATGCAAACTCATCTGAGTGGCGGACAGGAGGAGCAAAACA	5726
QY	5497	GGGCCAGGTCACTGTCTTTGGGCGCGGCTGCGGGGTGCACACCGAAAAATCTCTATCTCT	5556
DB	5727	AGGCAATGTTTCGGCTTTTGAAGAGGCGCTGAGGGGCCCAACACAGAAACCTACTCTCCAT	5786

QY	5557	GGACTTTGATCGGTAAACCCACAGAGAGAGATCTATGATGACCAACGCGAAGTTTACCCCT	5616
Db	5787	AGATTTTGTATCATATAACCCGCAACAGAAAGATCTATGATGACCATCGAAAAATTACCCCT	5846
QY	5617	TCGGATTCTGTACCAACAGCGGGGGCCAGCGCTCTGGT'CAACCAGCAGCAGGCTGAA	5676
Db	5847	TCGAATCTTTATGACCAGACTGGCGACCCATTCTGTGGTCTCTCTAAGCAGATATAA	5906
QY	5677	TGGTGTCAACGTGACATACCTCCCTCGTGGGGTTACATTTGCTGGCATCCAGAGGGGCATCAT	5736
Db	5907	TGAAGTGAACATCACATATTTACCTTCGGGATTTGGTGACGCTTTATTTCAAAAGAGGAACGTG	5966
QY	5737	GTCTGAAGAAGATGAATACGACACAGCGGGCGCGCATCACATCCAGGATCTTCCCTGATGG	5796
Db	5967	GAATGAATAAATGAATATGACCCAGAGTGGAAAAATTTATTTCAAGAACTTGGGCTGATGG	6026
QY	5797	GAAGACATGGAGCTACACATATTAGAGAAGTCCATGGTCTGCTACTACACAGCCAGAG	5856
Db	6027	GAATAATTGGAGCTATACCTACTTTAGAAAAATCTGTGATGCTTCTCCTCACAGCCAGCG	6086
QY	5857	GCAGTATATCTTTGAGTTTGACAGAATGACCGCTCTCTTGTGACGATGCCCCAACGT	5916
Db	6087	CGGTTACATCTTTGAGTATGACCAATCAGATTCGCTGCTGTAGTTACCATGCTGTAGCAT	6146
QY	5917	GGCGGGCAGACACTAGAGACCATCCCTCAGTGGGTACTACAGAAACATCTTATCAGCC	5976
Db	6147	GGTGGCCACAGCTTACAAACCATGCTTTCAGTGGGCTACTACGTAATATCTACACCCC	6206
QY	5977	CCCTGAGGGCAATGCCCTCAGTCAATACAGGACTTCACTGAGGATGGGCACTCCTTTCAAC	6036
Db	6207	ACCGGACAGTAGCACTCTTTTATCCAAAGACTATAGTCGAGATGGCCGATTTGCTACAGAC	6266
QY	6037	CTTCTACTGGCACTGGCGGAGGTGATATACAAGTATGGCAAACTGTCAAGCTGGC	6096
Db	6267	CCTGCATCTGGGACAGGGCGAGTCTTATCAAGTATCACCAAGCAAGCAAGGCTTTC	6326
QY	6097	AGAGACGCTCTATGACACCAAGGT'CAGTTT'CACTATGACGAGACGGCAGCGATGCT	6156
Db	6327	TGAGGTTCTCTATGATACCACTCAGGTCACATTAACATATGAAGAGTCTTCTGAGTGAT	6386
QY	6157	GAAGACCATCAACCTACAGAATGAGGCTTCACTGCACCATCCGTACCGTCAGATTGG	6216
Db	6387	TAAGACAATACACTGATGATGACGGATTCATCTGCACAATCAGATACAGGCAACACAGG	6446
QY	6217	GCCCTCATTTGACCGCAGATCTTCCGTTTCACTGAGGAAGGATGGTCAACGCCCGTTT	6276
Db	6447	ACCTCTTTTGGAGCCAGATTTTCAGATTCAGTGAAGAAGGCTTTGTGAATGCACGGTT	6506
QY	6277	TGACTACAATATGACAAACAGTTCGGGTGACACGATGACAGGTGTGATCAACAGAC	6336
Db	6507	CGACTACAGCT---ACAAACAATTTCCGAGT'CAACAGCTGCAAGCTGTATCAATGA AAC	6563
QY	6337	CCCACTGCCAATGTATCTCTATCGCTATGATGATGTGAGGCAAGACAGACAGTTTGG	6396
Db	6564	CCCTTGGCTATAGATCTTTACCGATATGTTGATGTCTCTGGCAGAAACAGAGCAGTTTG	6623
QY	6397	GAAGTTGGTGTCTATTTACTATGACATTAACAGATCATCACACAGCTGTCTATGACCCA	6456
Db	6624	AAAATT'CAAGTGAATTAATTACGATTTAAATCAGGT'CAATACTACTACAGTGATGAACA	6683
QY	6457	CACCAAGCATTTTGATGCATATCGCAGGATGAAGGAGTGCAGTATGAGATCTTCGCTC	6516
Db	6684	CACCAAAATCTTCAGTGGCCAAATGGAACAGT'CATTTGAAGTCCAAATATGA AATCTTAAGGC	6743
QY	6517	GCTCATGTACTGGATGACCGTCCAGTATGATTAACATGGGGCGAGTAGTGAAGAAG-GAGC	6575
Db	6744	AATTGCGCTACTGGATGACCATTCAAATATGATATGTTGGCCGACATGGTAAATATGTGACA	6803
QY	6576	TGAAGTAGGACCTTACGCCAATACCACTCGCTACTCTTATGATGATGATGCTGACGGCC	6635
Db	6804	TAAGGFTAGAGTAGATGGCAATATAAACAGGFACTTCTATGAATACGATGCTGTATGGCC	6863
QY	6636	AGTGTGACAGCAGTCTCCATCAATGAACAAGCACTCTGGGGCTACAGCTACGCTCTCAATG	6695

Db	6864	AACTTCAGACGTGTTCTGTAAATGACAAAAACCCAGTGGCGGTTTAGTTACGATCTCAATG	6923
Qy	6696	GGAACTCGCACTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCACTACGGTATGACA	6755
Db	6924	GAACAATCAACCTCTTAAGCCATGGGAAGAGTGTCTGTCTTACTCTCTCCGATATGACC	6983
Qy	6756	TCCGCAACCGCATCACTCGGCTGGTGACGTGCAATACAAGATGGATGAGGTGGCTTCC	6815
Db	6984	TCCGAGACCGCATCACCAGATTAGGAGAAATTCAGTATAAATGGATGAAGATGGGCTTTC	7043
Qy	6816	TGAGGAGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCCTGCTCATCAAGGCCT	6875
Db	7044	TGAGGCAGAGGGGAATGATATTTTGAATATTAATCTAAATGGCCTGCTCCAGAAAGCCT	7103
Qy	6876	ACAACCGGGTGGCAGCTGGAGTGTCAAGTACCCTACGATGCGCTGGGCGCGCGGTGT	6935
Db	7104	ACAATAGGCTTCTGGCTGGACTGTGCAGTATTACTATGATGGCTTGGCGCAGCTGTGC	7163
Qy	6936	CCAGCAAGAGAGCCACAGCCACCACTGCAGTTCCTATGACAGACCTGACCAACCCCA	6995
Db	7164	CGAGTAAGTCCAGCCTAGGCGCAGACCTTCAGTTCCTTTATGACAGACCTTACCAACCCCA	7223
Qy	6996	CCAAGTCAACCCACTGTACAACACATCCAGCTCTGAGATCACTCCCTCTACTACGACT	7055
Db	7224	TAAGAGTTACTCATTTGTACAACACACAAGCTCGAGATTACATCTCTGATTATGATC	7283
Qy	7056	TGCAAGACACCTCTTTGCCATGGAGCTGAGCTGTGTGATGAGTTTTCATGAGCTTGTG	7115
Db	7284	TCCAAGGTCACTTAITGCCATGGAGTTAAGCAGTGTGTAAGAATATTTGTAGCCTGTG	7343
Qy	7116	ACAACATCGGACCCCTCTTGCTCTTTAGTGGAAACAGGTTTGATCATCAAGCAAAATCC	7175
Db	7344	ATAATACAGTACCCCACTAGCTGTGTTGACGAGCCAGGTGAGGTCATTAAGGAGTAC	7403
Qy	7176	TGTACACAGGCTATGGGAGATCTACATGGATACAAACCCCACTTTGAGATCATCATAG	7235
Db	7404	TATACACACTTATGGCGATCTATCATGACACATTACCTGACTTTTCAGGTCAITATTG	7463
Qy	7236	GCTACCATGGTGGCTCTATGATCCACTCACCAAGCTGTGCCATGGCGCGGAGATT	7295
Db	7464	GTTTTTCATGAGGACTCTATGATTTTCTTAAATTAGTGCACTCTGGGCGAAGGATT	7523
Qy	7296	ATGATGTGCTGGCCGAGCTGGACTAGCCAGACCCAGAGCTGTGGAAGCACTTAGTA	7355
Db	7524	ATGATGTTGTTGTCGGCAGATGGACAACGCCCTAATCATCATATGGAACAGTT--GA	7580
Qy	7356	GCACCAAGCATGCTGCTTTTAATCTCTATGTTCAAACAACAACCCCATCAGCACT	7415
Db	7581	ACCTCCTTCCTAACCAATTCACCTCTACTCCTTTGAAAATACTACCCAGTTGGCAAAA	7640
Qy	7416	CCCAGGACATCAAGTGCTTCATGACAGATGTTAACAGCTGCTGCTCACCTTTGGATCC	7475
Db	7641	TTCAAGATGTGCAAGATATACACAGACATCAGAAGTTGTTGGAGCTATTTGGTTTCC	7700
Qy	7476	AGCTACACAACGTGATCCCTGGTATTCCAAACACAGACATGGATGCCATGGAACCTCTCT	7535
Db	7701	AAATTACAAATGTACTCTCGATTTCCCAACCTGAAATTAGAAAAATTAGAAATTAACTT	7760
Qy	7536	ACGAGCTCATCCACACACAGATGAAAACGCGAGGAGTGGGCAACAGCAAGTCTATCTCG	7595
Db	7761	ACGAGCTTCTACGGCTTCAGACAAAACCTCAAGAGTGGGATCTCTGGAAGACTACTCTGG	7820
Qy	7596	GGGTACAGTGTGAAGTACAGAGCAGCTCAAGGCTTTGTCACTTTAGAACGGTTTGACC	7655
Db	7821	GCAATTCAGTGTGAACCTCCAGAAAACAGCTCAGGAATTTCAATTCCTTGGACCACTAGCTA	7880
Qy	7656	AGCTCTATGGCTCCACAATCACAGCTGCCAGCAGGCTCCAAAGACCAAGAGTTTGCAT	7715
Db	7881	TGACTCCCCGATACAATGATGGAACGTGCTTTGAAGGAGGAGCAACCAAGTTTGTCTG	7940
Qy	7716	CCAGCGGCTCACTTTTGGCAAGGGGCTCAAGTTTGCTTTGAAGGATGGCCAGTGACCA	7775

Db	7941	CTGTCCCTCTGTGTTTTTGGGAAAGGTATAAAAAATTTGCCATCAAGGATGGCATAGTAACAG	8000
Qy	7776	CAGACATCATCAGTGTGGCCCAATCAGAGATGGGCGAAGGGTTGCTGCGCATCTTGAACCAATG	7835
Db	8001	CTGATATTAAGGAGTAGCCAAATGAAGATAGCAGGCGGCTTGCTGCCATTTCAATAATG	8060
Qy	7836	CCCACTACCTTAGAGAACCTGCACCTTCCACATTTGATGGGGTGGATACCCATTACTTTCTGTGA	7895
Db	8061	CCCAATTACCTTGGAAAAACCTTACATTTTACCATAGAGGGAGGGACACTCATCTCTTCAATTA	8120
Qy	7896	AACGAGACCTTTCAGAAAGGTGACCTGGCCCATCTCGGCCCTCAGTGGGGGCGCGGAACCC	7955
Db	8121	AGCTTGGGTCTCTCGAGGAAGACCTGGTGTCTCATCGGTAACTATGGGGGAGCGCGATTTC	8180
Qy	7956	TGGAGAAATGGGGTCAACGTCACCTGTGTCTCCAGATCAACACAGTACTTAAATGGCAGGACTA	8015
Db	8181	TGGAGAAATGGTGTCAATGTCTCACTGTGTCCAGATGACTTCTGTGTTGAATGGGAGGACTA	8240
Qy	8016	GAGCTTACACAGACATCCAGCTCCAGTACCGGGGCACTGTGCTTGAACACACAGCTTACCGGA	8075
Db	8241	GACGGTTTGAGATATTTTCAGCTCCAGCATGGAGCCCTGTGCTTCAACATCCGGTATCGGA	8300
Qy	8076	CAACGTTTGGATAGGAGGAAGCGACCGGTCTCTGAGCTTGGCCCGGCGCAGAGACCGTTCGCC	8135
Db	8301	CAACTGTCTGAGAGGAAAGAANAATCACGTGTGTGGAGATTGCGCAGACGCGCGCAGTGGCCC	8360
Qy	8136	AAGCGTGGGCCCGCGAGCAGACAGACTGCGGGAAGGGAGGAGGCGCTTCGGGCGCTTGGGA	8195
Db	8361	AGGCTTGACTAAGGAACAAGAAGGCTGCAAGAGGGGAAGAGGGGATTAGGGCATGGA	8420
Qy	8196	CAGAGGGGGAAGCAGCAGCGTGTCTGAGCAAGGGCGGGTGCAGAGGCTACGACGGCTTTT	8255
Db	8421	CAGAAGGGGAAAAGCAGCAGCTTTTGAGCACCTGGGCGGGTACAAGGTTACGATGGGTATT	8480
Qy	8256	TCGTGATCTCTGTGCGACGCTACCCAGAACTGTCCAGACGCGCCCAACAACATCCACTTCA	8315
Db	8481	TTGTTTGTCTGTTTGGACAGTATTTAGAACTTCTGACGTGCCCAATAATATTCACTTTA	8540
Qy	8316	TGACACAGCGCAGATGGCGGAGGTGCACAGAGA	8350
Db	8541	TGACACAGCGCAATAAGCAGGAGGTAACAAAAA	8575
RESULT 13			
ABK92230			
ID	ABK92230 standard; DNA; 12879 BP.		
XX			
AC	ABK92230;		
XX			
DT	15-AUG-2002 (first entry)		
XX			
DE	Prostate cancer-associated DNA sequence #116.		
XX			
KW	Prostate cancer; prostate tumour tissue; human; mammal; cytostatic;		
XX			
KW	gene therapy; gene; ds.		
XX			
OS	Mammalia.		
XX			
PN	WO200230268-A2.		
XX			
PD	18-APR-2002.		
XX			
PF	12-OCT-2001; 2001WO-US032045.		
XX			
PR	13-OCT-2000; 2000US-00687576.		
XX			
PR	08-DEC-2000; 2000US-00733288.		
XX			
PR	08-DEC-2000; 2000US-00733742.		
XX			
PR	24-JAN-2001; 2001US-0263957P.		
XX			
PR	16-MAR-2001; 2001US-0276791P.		
XX			
PR	16-MAR-2001; 2001US-0276888P.		
XX			
PR	06-APR-2001; 2001US-0281922P.		
XX			
PR	24-APR-2001; 2001US-0286214P.		
XX			
PR	30-APR-2001; 2001US-00847046.		

RESULT 13

ABK92230	ABK92230 standard; DNA; 12879 BP.
XX	
AC	ABK92230;
XX	
DT	15-AUG-2002 (first entry)
XX	
DE	Prostate cancer-associated DNA seq
XX	
KW	Prostate cancer; prostate tumour
XX	
KW	gene therapy; gene; ds.
XX	
OS	Mammalia.
XX	
PN	WO200230268-A2.
XX	
PD	18-APR-2002.
XX	
PF	12-OCT-2001; 2001WO-US032045.
XX	
PR	13-OCT-2000; 2000US-00687576.
PR	08-DEC-2000; 2000US-C0733288.
PR	08-DEC-2000; 2000US-C0733742.
PR	24-JAN-2001; 2001US-0263957P.
PR	16-MAR-2001; 2001US-0276751P.
PR	16-MAR-2001; 2001US-0276888P.
PR	06-APR-2001; 2001US-0281922P.
PR	24-APR-2001; 2001US-0286214P.
PR	30-APR-2001; 2001US-00847046.

PR 04-MAY-2001; 2001US-0289589P.
 PA (EOSB-) EOS BIOTECHNOLOGY INC.
 PI Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
 XX WPI: 2002-471335/50.
 DR P-PSDB; ABG61913.
 XX
 PT Detecting a prostate cancer-associated transcript in a cell in a patient,
 PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
 PT by determining if prostate cancer-associated genes are expressed in a
 PT prostate tissue.
 XX
 PS Claim 22; Page 394-397; 436pp; English.
 XX
 CC The present invention relates to methods of detecting a prostate cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with prostate cancer-
 CC associated polynucleotides (designated PC genes) that selectively
 CC hybridize to a sequence that is at least 80% identical to them. The
 CC prostate cancer-associated polynucleotide sequences are differentially
 CC expressed in prostate tumour tissue or in prostate cancer and are derived
 CC from the tissues of various organisms such as humans or other mammals
 CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
 CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
 CC associated genes are useful for diagnosing or treating prostate cancer,
 CC as well as for identifying modulators of prostate cancer or agents that
 CC inhibit prostate cancer. The nucleic acid sequences are particularly
 CC useful in gene therapy, as a vaccine or in antisense applications.
 CC ABK92115-ABK92263 represent prostate cancer-associated polynucleotide
 CC sequences
 XX
 SQ Sequence 12879 BP; 3852 A; 2674 C; 2768 G; 3585 T; 0 U; 0 Other;
 Query Match 39.0%; Score 3258; DB 6; Length 12879;
 Best Local Similarity 65.5%; Pred. No. 0;
 Matches 4893; Conservative 0; Mismatches 2515; Indels 66; Gaps 6;
 QY 880 CGGCACTTCCTCTCAAGCTGGA---GGCACTCCCGCTCTCTGCAACCATCACC 936
 DB 838 CAGGCAATTCCTGTTCAAAATGATCTGGTTCCTCTGCGATCTTCAAGTCAGCCAGTCA 937
 QY 937 AGGTAACCACTGACGTCACAGCAGTGTACTCTCTCCCGCCCGACCCCTGCGCCGCGAG 996
 DB 998 GAATACCTCTGACATCCAAATACCGTGTACTCGCCCTCTCCAGGCTCTTCTCTGAAG 957
 QY 997 CACCTTGGCCCGCGCGCTTTAACTCAAGAGCCCTCAAGTACTGTAACTGGAAGTG 1056
 DB 958 CACCTTTTCCCGACTGCTCTTAACTTTAAACAACTTACAGGTGCTGCACTGGAAGTG 1017
 QY 1057 CGCAGCCCTGAGCCCACTGTCATCTCAGCCACTCTGTCATCTCTGTCGACATCTTGT 1116
 DB 1018 CACAGCATTTGAGCGCCATGCAATCAGTGACTTTGGCTTTGTTACTAGCCTATGTGAT 1077
 QY 1117 GGCCATGCACTGTTGGCTTAACTGGACCTGCGACCGCATGAGGGGAGATGATGA 1176
 DB 1078 TGCAGTGCAATTTGTTGGCTGCTGACTTGGCAGTTGCAACAGTTTGAAGAGAGAGTGTATGC 1137
 QY 1177 GATCAGCGAGGACACAGCAGCAGTGTGGCTGTGTGCGCAACCGAGCTCTCCCTATACCCCTC 1236
 DB 1138 AATGGAGTTAGCAAGGGACAGGGGACCGAGTCCATGGACACTTACTTCTCCAAAT 1197
 QY 1237 AGGGGCACTGGCTTAGAGACCCCTGACAGGAAAGGCAAGAACACAGAGGAAAGCC 1296
 DB 1198 TGCAGGAAAAGTTTCTGATAAATCAGAGAAAAG----- 1232
 QY 1297 CAGTAGTTTCTTCCAGAGACAGTTTCATAGATTCTGGAGAAATTCATGTGGAGAGCG 1356
 DB 1233 -----TGTTTCAGAAAGGACGGGCGATAGACACTGGAGAAAGTTGACATTGGTGCA 1284
 QY 1357 AGCCTCCAGAAAGTTCCTCTGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCC 1416

DB 1285 GGTATGTCAGACCAATTCACCTGGTTTATTCCTGGGTTTCCAGATTACTATCCACCATCC 1344
 QY 1417 TGTGATCTGAATTCATATGTCTCTGGAAGGAGCCCTGGTTGGCATTTATGGCAG 1476
 DB 1345 AATATATCTGAAGTTCAATATTTCTTAGCCAGAGACTCTCTGCTGGAAATTTATGGCAG 1404
 QY 1477 AAAAGCCCTCCCTCTCTCACATACACAGTTTGTAGCTTTGTGGAGCTGTGGATGGCAGAG 1536
 DB 1405 AAGAAACATTCACCTCACATACATCTCAGTTTGTATTTGTATAAACTAAATGGATGGCAACA 1464
 QY 1537 GTCCTTAACCCAGGAGGCGGAGCCTAGAGGGGACCCGCGCCAGTCTCGGGAGACTGT 1596
 DB 1465 GCTGTCAAGCAGGACTCAAGGGCTCTGATATACAGACACTCCCTCCGAACTGAT 1524
 QY 1597 GCCCCCTCCAGCCATGAGACAGGCTTCATCAGTATTTGGATTTCAGGAATCTGCACTT 1656
 DB 1525 CTTAACTTCGCTTCAGGAGACAGGTTTCATAGAGTATATGATCAAGGACCTTGTATCT 1584
 QY 1657 GCTTTTACATGACGGAAGAGTCAAGAGTGGTTTCTTTCTCACCACCTGCCATGA 1716
 DB 1585 GCGGTTTACATGATGGAAGGATGAGCAAGTATTCGTGTAACTACAGCAATTTGA 1644
 QY 1717 GTGCTGTGATAACTGCCCCAGCACTGTCTATGGCAATGTGTGACTGCTCTGGACCTG 1776
 DB 1645 AATAATGATGACTGTTCACCAATTGCAATGGAAATGAGAGTGTATCTCTGGCCATTG 1704
 QY 1777 CCATGCTTCTGTGGTTTCTTGGGCCCGGACTGTGGCAGAGCCCTCTGCCCCGCTGCTG 1836
 DB 1705 TCATTGTTTCCAGGATTCCTTGGACCTGACTGTGCTAGAGATTCCTGCCCCCTGCTGTG 1764
 QY 1837 TAGCGGAATGCCAATACATGAAGGAGAGTGTGTGTCACAGTGTGCTGCGGAAGCGC 1896
 DB 1765 TGTGGAATGGAGAAATGAGAGACACTGTGTCTGCCGATGCTGCGAAGGGGCC 1824
 QY 1897 TGAGTGCATGTGCCCAACCAACAGTGTATCGATGTGCTGTCAGCAACCAATGCACTG 1956
 DB 1825 AGAGTGTGACGTTCCGGAAGAACAAATGATTCATGATCCAACTGCTTTGGCCACGCACTG 1884
 QY 1957 CATCAGCGGACCTTCATCTGCAACCTGCTGCTAGCAAGGGGAGAGCTGTGAGGAGTGA 2016
 DB 1885 CATCATGGAGTGTGATCTGTGTCAGGATACAAAGGAGAAATATGCGAGGAGAGGA 1944
 QY 2017 CTGATGCAACCCCAATGTTTCAGGCCGGGTGTCTGGTGTAGAGGCGAATGCCATGCTT 2076
 DB 1945 CTGCTAGACCAATGTGTTCCAACTGCTGCTGTAAGAGAGATGTCTACTGTT 2004
 QY 2077 TGTGGGATGGGAGGACCAACTGCGAGACCCCGAGGCCACATGCTTTAGACAGTGTTC 2136
 DB 2005 TACTGGCTGGGAGAGTTAACTGTGAAACACCACTTCTCTGTATGTCAAGAGAGTGTCT 2054
 QY 2137 AGGCCACGAACTTCCTCCCGGACACCGGGCTTTGCACTGTGTCACCAACCTGCACTGG 2196
 DB 2065 AGGACACGAACTTTCTTCTGGAGCTGGAGTATGACAGTGTGATCCCAAGTACAGAG 2124
 QY 2197 ACAAGCTGTTCTATGAGATCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2256
 DB 2125 ATCTGACTGTCTCAACAGAGCTGTGATCCATGAGTGTGGTAGCCATGAGTGTCTCTCAAG 2184
 QY 2257 GGGCACCTGCGCTGCGAGGATGCTGATGGGGGAGCCCTGCGACAGCGGGCTGCCA 2316
 DB 2185 AGGAATTTCCAGTGTGAGAGGCTGGGTAGGACCAACATGTGAGGAACGCTCTCTGCA 2244
 QY 2317 CCCGCTGTGCGAGCATGGGACCTGCGCGAGCGCAAGTGGAGTGGAGTGGAGTGGAGTGG 2376
 DB 2245 TTCTCATTTGACTGAGCATGCCCAATGCAAGATGGAATAATGTGAGTGTAGCCCTGATG 2304
 QY 2377 GAATGGCGAACACTGCCACCTGCTCACTATCTGGATAGGAGTGTAAAGAGGGTTCGCC 2436
 DB 2305 GGAGGGGACCACTGSCAAATTTGCTCACTACTTAGAT---GCTGTCGAGATGGCTGCC 2361
 QY 2437 TGGGTTGTGCAATGGCAACCGCAGATGTACTGTAGACCTGAATGGTTGGCAGCTGCTG 2496
 DB 2362 AGGCTCTGCTTTGGAAATGACAGATGTACCTGGATCAAAATGGTTGGCACTGTGTGTG 2421

Db 4561 CTTAGCAGTGTGCTGATGAAACCCCTCTATGTGCGAGACCTCGAAATGTGCAATTCG 4620
Qy 4717 GTTATCCGGAAGAACAGACCTTTCTCAACACCCAGAAACATGTATGAGTGTCTTCAC 4776
Db 4621 TACCATCAGCAGGAACCAAGCCCACTGAAATGACATGAACATTTATGAGATTCGTTCA 4680
Qy 4777 AATTGACAGGAGCTCTATCTGTGTTGATACACCGGCAAGCACTGTACACCCAAAGCCT 4836
Db 4681 CGCTGATCAGGAATGTACCAATTCAGTAAATGGAACCCACCTACACACCTTGACCT 4740
Qy 4837 GCCACAGGAGACTACCTGTACAACTTCACTACCTACCTGCGGAGCGGACATCACTCAT 4896
Db 4741 GATTAACAGGAGACTATGTTATTAACCTTCACTCAATTCGAAGTGACTTGGCGCGAT 4800
Qy 4897 CACAGACAAACATGCAACATGTTAAATGTCGCGGAGACTCTACTGGGATGCCCTCTG 4956
Db 4801 TACCAGACGAATGGCAATTCAGTGCACATTCGCGGTGATGCGAGCGGAATGCCGTATG 4860
Qy 4957 GCTGCTGCTCCAGATGGCCAGGTGTACTGGGTGACCATGGGACCAACAGTGCATCAA 5016
Db 4861 GCTTGTGCTGCTGGCGACAAAGTATATCTGCTGACTATAAGCAGCAATGGAGTCTGAA 4920
Qy 5017 GAGTGTGACACACAAAGGACAGAGTTGGCCATGATGACATACATGGCAATTCGGCCT 5076
Db 4921 AAGAGTGTGACGCCAAGGCTATAATCCGGCTTAAATGACCTATCCAGGAACACAGGGCT 4980
Qy 5077 TCTGCCAACCAAAAGCAATGAACCGGATGGAACAACTTTATGAGTACGACAGCTTTGG 5136
Db 4981 TCTGCTACCAAAAGTAAAGCAATGATGAGCAACCGTTATGAGTATGACCCGAGGG 5040
Qy 5137 CCGCTGACAAATGTGACTTCCCTACTGGCCAGTGAGCAGTTTCGGAAGTATACAGA 5196
Db 5041 ACACCTGACCAATCAACAGTTTCCCACTGGAGAGGTGAGCAGCTTCCACAGTACCTGGA 5100
Qy 5197 CAGTTCAGTGATGTCAGGTGAGAGCTCCAGCAAGATGATGTCACCATACCAACCAA 5256
Db 5101 GAAGCTGACAAAGTGAAGTATGATTTCCAAACCGTGAATGCTCTATGTCAACCAA 5160
Qy 5257 CCTGTGCTCAGGCGCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTA 5316
Db 5161 CTTGACGGCACTAGTACCATATATATTTAAACAAAGAAATACTCAAAGTACCTATCG 5220
Qy 5317 CATCGGGCGAGTGGCTTCTGCGGCTGCTGCTGGCCAAAGGAGTGGCGGTGCA 5376
Db 5221 GGTGAATCCAGATGTTCCCTGGGTGTCATTTTGGCCAGGGATGGAGATGGGCTCAG 5280
Qy 5377 GACTGAGCCCACTGCTGGCTGGCACCGTCAACCCCGGCAAGAGGAATGTCA 5436
Db 5281 CTCAGAGCCCACTCTGGCAGGGGAGTCAACCTTACCTCGGCAATGCAACATCTC 5340
Qy 5437 GCTGCCATCGAACCGGCTCAACCTGTGGTGGAGTGGCGGCGCAAGAGAGGCTCG 5496
Db 5341 ATTGCCCGGAGAGACAAATCAACCTTCATCGAGTGGCGGAGAGGAGGAGGAGGAG 5400
Qy 5497 GGGCCAGTCACTGCTTTTGGGCGCGGCTGGGCTGCAACCGGAAATCTCTATCTCT 5556
Db 5401 AGGCAATGTTTGGCTTTTGAAGAGGCTGAGGGGCCCAACAGAAACCTACTCTCAT 5460
Qy 5557 GGACTTTGATCGCGTAACAGCAGAGAGATCTATGATGACACCGCAAGTTCACCT 5616
Db 5461 AGATTTTGTATATATACCGGCAAGGAAAGATCTATGATGACCATCGAAATTCACCT 5520
Qy 5617 TCGATTTCTGTAGCAGCGGGGGCGGCGGCTGCTGTCACCCAGGAGGAGGCTGAA 5676
Db 5521 TCGAATTTCTTATGACCACTGGGCGACCATTTCTGTGCTCTCTGTAGCAGATATAA 5580
Qy 5677 TGTGTCAACGTCATCTCCCTGGGGTTACATTTGCTGGCATCCAGGGGCTCAT 5736
Db 5581 TGAAGTGAACATCACATATTCACCTCGGGATTTGAGCTTTATTTCAAAGGAGAACGTG 5640
Qy 5737 GTCTGAAGATGGAATACACAGCGGCGGCGGCATCATCAGGATCTTCTGCTGATCG 5796

Db 5641 GAATGAAAAATGGAAATATGACCAGAGTGGGAAAAATTTATTTCAAGAACTTGGGCTGATGG 5700
Qy 5797 GAAAGACATGGAGCTACACATCTACTAGAGAAAGTCCATGCTGTCTACTACACAGCCAGAG 5856
Db 5701 GAAAAATTTGGAGCTATACCTACTTAGAAAAATCTGTGATGCTTCTCTACACAGCCAGCG 5760
Qy 5857 GCAGTATATCTTTGAGTTCGACAAAGATGACCGCTCTCTTCTGTGACGATGCCCAACGT 5916
Db 5761 GGTATACATCTTTGAGTATGACCAATCAGATTCCTGCTGTCTAGTACCATGCTAGCAT 5820
Qy 5917 GCGCGGAGACACATAGAGACATCCGCTCAGTGGGTACTACAGAAACATCTATACAGCC 5976
Db 5821 GGTGGCCACAGCTTTACAAACCATGTTTCACTGGGCTACTACCGTAAATATCTACACCC 5880
Qy 5977 CCTCAGGCGCAATGCTCAGTCATACAGAGCTTCACTGAGGATGGGACCTCTTTCACAC 6036
Db 5881 ACCGACAGTACGACTTCTTTATCCAAAGACTATAGTCGAGATGGCGATGCTACAGAC 5940
Qy 6037 CTTTACCTGGGCACTGCGCCAGGGTGATATACAAGTATGGCAAACTGTCAAAGCTGGC 6096
Db 5941 CCTGCATCTGGGACAGGCGCAGAGTCTTATACAAGTACACCAAGCAAGCAGGCTTTC 6000
Qy 6097 AGAGAGCTCTATGACACCAACCAAGGTCACTTCACTATGACAGAGCGGAGGATGCT 6156
Db 6001 TGAGGTTCTCTATGATACCACTCAGGTCACTTAACATATGAAGAGTCTTCTGGAGTGA 6060
Qy 6157 GAAAGACCATCAACCTACAGAAATGAGGCTTCACTGCAACCATCCGCTACGATGTTG 6216
Db 6061 TAAGACATACACCTGTATGATGACGAGTTCATCTGCACAAATCAGATACAGGCAACAG 6120
Qy 6217 GCCCTGATGACGACAGATCTTCCGCTTCACTGAGGAAGGATGCTCAACGCCCTGTT 6276
Db 6121 ACCTTTATTGGCGCCAGATTTTCAATTCAGTGAAGAGGCTTGTGAATGACGCTT 6180
Qy 6277 TGACTTACAACTATGACAAACAGCTTCCGGGTGACCAAGATGAGGCTGTGATCAACAGAG 6336
Db 6181 CGACTACAGCT--ACAACAATTTCCGAGTCAAGCATGCAAGCTGTATCAATGAAC 6327
Qy 6337 CCCACTGCCAATGATCTATCGCTATGATGATGTGTCAGGCAAGACAGAGAGTGTGG 6396
Db 6238 CCGTTTGGCTATAGATCTTTACCGATATGTTGATGCTCTGGCAGAAACAGAGAGTGTGG 6297
Qy 6397 GAAGTTTGGTCTCATTTACTATGATTAACAGATCATCACACAGCTGTCTATGACCCA 6456
Db 6298 AAAATTCAGTGAATTAATTAAGATTTAATCAGGTCACTAACTACTACAGTATGAACA 6357
Qy 6457 CACCAAGCATTTTGAATGATGCAATGCGAGGATGAAGAAAGTCAGTATGAGATCTTCCGCTC 6516
Db 6358 CACCAAAATCTTCAGTGCATGCAAGTCAATGAAAGTCCAATATGAATCTTAAAGGC 6417
Qy 6517 GCTCATGTACTGAGTACCGCTCAGTATGATAACATGGGGGAGTATGAAGAGAGCT 6576
Db 6418 AATTGCCCTACTGATGACCAATTCATATGATTAATGTGGCGGACATGTTAATATGTGAT 6477
Qy 6577 GAAAGTAGGACCTTACGCCAATACCACTCGCTACTCTCTATGATGATGATGTCAGCGCCA 6636
Db 6478 AAGGTAGGAGTATGCGCAATATAACAAGTACTTCTATGAATACGATGCTGATGGCA 6537
Qy 6637 GCTGACAGAGCTCTCAATGACAGGCACTCTGGGGCTACAGCTACGACCTCAATGG 6596
Db 6538 ACTTCAGACTGTTTCTGTAAATGACAAAACCCAGTGGCGTTATAGTTACGATCTGATGG 6597
Qy 6697 GAACTGCACATCTTCTGAGCCCTGGGAAACAGTGCAGGCTCACACCACTACGGTATGACAT 6756
Db 6598 AGACATCAACCTTTAAGCCATGGAAGAGAGTCTGCTTCTTCTCTCCGATATGACCT 6857
Qy 6757 CCGGACCGCATCACTCGGCTGGGTGAGTGCATACAGATGAGATGAGATGCTTCTCT 6816
Db 6658 CCGAGACCGCATCAACAGATTTAGGAGAAATTCAGTATAAAATGGAATGAAGATGGCTTCT 6717
Qy 6817 GAGCAGCGGGCGGCTGATCTTTGAGTACAACTCAGCTGGCTGCTCATCAGGCTA 6876
Db 6718 GAGCAGAGGGGAAATGATATTTTGAATATTAATTTTGAATGCTGTCGAGAAAGCCTA 6777

QY 6877 CAACCGGCTGGCAGCTGGAGTGTACAGTACCGTACAGTGGCTGGGGGGCGCGTGTG 6936
 Db 6778 CAATAAGGCTTCTGGCTGGAGTGTGAGTATTAATATGATGGCTTGGGGGAGCTGTGCG 6837
 QY 6937 CAGCAAGAGCAGCAGCAGCAGCAGCAGTCTTCTATGCGAGCTGACCAACCCAC 6996
 Db 6938 GAGTAAGTCAGCAGCAGCAGCAGCAGTCTTCTATGCGAGCTGACCAACCCAC 6997
 QY 6997 CAAGGTCAACCCAGCTGACCAACCACTCCAGTCTGAGATCACCTCCCTCTACTACGACTT 7056
 Db 6998 AAGAGTTACTCATTTGTACAAACCAACAAGCTCGGAGATTACATCTCTGTATTATGATCT 6957
 QY 7057 GCAGGAGCACTCTTTCCGATGGAGCTGAGAGTGGTGTGATGATTTTACATAGCTGTGA 7116
 Db 6958 CCAAGGTCACTTATTCGATGGAGTGAAGCAGTGTGTAAGATATTTATAGCCCTGTGA 7017
 QY 7117 CAACATCGGAGCCCTCTGCTGCTTTAGTGAACAGGTTTGTATGATCAAGCAAAATCCT 7176
 Db 7018 TAATAAGGTACCCCACTAGCTGTGTTACAGCCGAGGTGAGTCTATAAGAGATACT 7077
 QY 7177 GTACAGCCATATGGGAGATCTATATGATGATCAACCCCACTTTGAGATCATATAGG 7236
 Db 7078 ATACACACCTTATGGGATATCTATCATGACACTTACCTGACTTTGAGTCTATAATTGG 7137
 QY 7237 CTACATGCTGGCTCTATGATCACTCAACAGCTTGTCCATGCGCGGCGAGATTA 7296
 Db 7138 TTTTCATGGAGGACTCTATGATTTCTTAAATAGTGCACCTGGGGGCAAGGATTA 7197
 QY 7297 TGATGCTGCTGGCGGAGCTGGAGTACGCTAGCCAGACAGAGCTGTGGAAGCACCTTAGTAG 7356
 Db 7198 TGATGTTGTTGCTGGCAGATGGACAACGCGCTATCATCATATGGAACAGTT--GAA 7254
 QY 7357 CAGCAAGCTGATGCTTTTATCTCTATATGTTCAAAACCAACACCCCATCAGCACTC 7416
 Db 7255 CCTCCTTCTTAACCACTTCACTCTACTCTCTTGAATAAATACTACCCAGTTGGCAAAAT 7314
 QY 7417 CCAGGACATCAAGTCTCTATGACAGATGTTAAAGCTGCTGCTCACCTTTGGATTCCA 7476
 Db 7315 TCAGATGTTGCAAGTATACCAAGACATCAGAGTGTGTTGAGGATATTTGGTTCCA 7374
 QY 7477 GCTACAAAGTGTATCCCTGTTATCCAAACCAAGATGATGATGATGATGATGATGATGAT 7536
 Db 7375 ATTACAAATGTACTACCTGATTTCCCAAACTGAAATTTAGAAATTTAGAAATTA 7434
 QY 7537 CGAGCTCATCACACAGATGAAAGCAGGAGTGGAGACACAGCAAGTCTATCTCTCGG 7596
 Db 7435 CGAGCTTCTAGGCTTCAAGCAAAACTCAAGAGTGGGATCCTGGAAGAGTATCTCTGGG 7494
 QY 7597 GGTACAGTGTGAAGTACAGAGCAGCTCAAGGCTTTGTCACTTTAGAAAGGTTTGACCA 7656
 Db 7495 CATTCAGTGTGAAGTCCAGAAACAGCTCAGGAATTTCAATTTCTTGACCAACTACCTAT 7554
 QY 7657 GCTCTATGGCTCCACATCACCAGTGGCAGAGCTCCAAAGACCAAGAGTTTGATC 7716
 Db 7555 GACTCCCGATACAAATGATGACGCTGCTTTGAAGAGGGAAGCAACCAAGTTTGTGTC 7614
 QY 7717 CAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGATGCGCGAGTGACCA 7776
 Db 7615 TGTCCCTCTCTTTTGGGAAGGTATAAATTTGCCATCAAGATGCGATAGTACAGC 7674
 QY 7777 AGACATCATAGTGTGCGCAATAGGATGGCGAGGTTGCTGCCATCTTGAACCATG 7836
 Db 7675 TGATATTAGAGTAGCAATGAAGATAGCAGCGGCTTGTGCCATTTCTCAATAATGC 7734
 QY 7837 CCAGTACTAGAGACCTGCACTTCACTATGATGGGTGGATACCCATTTCTGTGA 7896
 Db 7735 CCATTACTGGAACCTTACATTTTACCATAGAGGGAGGAGCACTCACTACTTCAATA 7794
 QY 7897 ACCAGGACCTTCAGAAAGTCACTGGCCATCTGGGCTCAGTGGGGGGGGGCAACCT 7956
 Db 7795 GCTTGGGTCTCTGGAGGAGACCTGTGCTCATCGGTAACACTGGGGGGGGGGGATTTCT 7854

QY 7957 GGAGAAATGGGTCAACGTCACTGTCTCCAGATCAACACAGTACTTAATGGCAGACTAG 8016
 Db 7855 GGAGAAATGGTGTCAATGTCACTGTCTCCAGATCACTTCTGTGTAATGGGAGACTAG 7914
 QY 8017 AGCTTACACAGACATCCAGCTCCAGTACGGGCACTGTGTGAACACACGCTACGGAC 8076
 Db 7915 ACGTTTTCAGATATTCAGCTCCAGCATGGAGCCCTGTCTTCAACATCCGGTATGGAC 7974
 QY 8077 AACGTTGGATGAGGAGAGGCAAGGCTCTCGAGCTGCGCCGCGCAGAGAGCGTGCACA 8136
 Db 7975 AACTGTGGAAGAGGAAAGAAATCACGTGTGGAGATTGCCAGACAGCGCAGTGGCCCA 8034
 QY 8137 ACGTGGGCCCCGAGCAGCAGAGCTCGGGAAGGGAGGAGGAGCCCTGCGGCTGGAC 8196
 Db 8035 GGCCTGGACTTAAGGAACAAAGAGGCTCAAGAGGGGGAAGAGGGATTAGGCGATGGAC 8094
 QY 8197 AGAGGGGAGAGAGCAGCAGGCTCTGAGCACAGGCGGGTGAAGGCTACGACGGCTTTT 8256
 Db 8095 AGAGGGGAGAGAGCAGCAGGCTTTTGGACACTGGGCGGGTACAGGTTACGATGGTATTT 8154
 QY 8257 COTGATCTCTGTGAGCAGTACCCAGAACTGTACAGACGCGCAACACATCCACTTCAT 8316
 Db 8155 TGTGTTGCTGTGTGAGCAGTATTTAGAACTTTTGCACAGTCCCAATAATATTCACTTTAT 8214
 QY 8317 GAGCAGAGCAGATGGCGCGGAGGTGACAGAGA 8350
 Db 8215 GAGCAGAGCGAAATAGGAGGAGTAAACAAA 8248

RESULT 14
 ACC72052

ID ACC72052 standard; DNA; 9695 BP.

AC ACC72052;

XX 08-JUL-2003 (first entry)

XX BCU0205B gene #SEQ ID 81.

XX Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;
 XX drug discovery; clinical medicine; forensic medicine; gene;
 XX chromosome 5q33.3; ds.

XX Homo sapiens.

XX WO2003029421-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031287.

XX 03-OCT-2001; 2001US-0326526P.

XX 14-MAY-2002; 2002US-00144194.

XX (ORIG-) ORIGENE TECHNOLOGIES INC.

XX Sun Z, Li X, Fan W, Kovacs KF, Jay G;

XX WPI: 2003-381623/36.

XX P-PSDB; ABR58318.

XX New isolated human differentially-regulated breast cancer polynucleotide
 XX and polypeptide, useful for diagnosing, staging, prognosticating,
 XX preventing and/or treating diseases and conditions relating to breast
 XX cancer.

XX Claim 2; SEQ ID NO 81; 127pp + Sequence Listing; English.

XX The invention relates to isolated polynucleotides which are
 XX differentially-regulated in breast cancer. The methods and compositions
 XX of the present invention are useful for detecting, diagnosing, staging,
 XX monitoring, prognosticating, preventing and/or treating diseases and
 XX conditions relating to breast cancer, and may be used in gene therapy or

QY 6383 ACAGAGCAGTTTGGAAAGTTTGGTGTCTATTTACTATGACATTAACAGATCATCACCACA 6442
Db GTGGAACACCTTGGTAAAGTTTGGAGTCTATCTATTATGACATCAACAGATCATCACCAC 6446
QY 6443 GCTGTATGACCCACACACAGCAATTTTGAATGCATATGCGAGGATGAAGAGTGCAGTAT 6502
Db GCCGTGATGACCCCTCAGCAACACTTCGACACCCATGGCGGATCAAGGAGGTCCTAGTAT 6506
QY 6503 GAGATCTTCGCTCGCTCATGTACTGTGATGACCGTCCAGTATGATACATGCGGCGAGTA 6562
Db GAGATGTTCCGTTCCCTCATGTACTGTGATGACCGTGCATATGACAGCATGGCGGGTG 6566
QY 6563 GTGAAGAGGAGTGAAGTAGGACCTTAGCCGCAATACCACTGCTACTCTCTATGATAT 6622
Db ATCAAGAGGAGCTTAAACTGGGCGCTATGCGCAATACCAAGAGTACACCTATGATAC 6626
QY 6623 GATGCTGACGGCAGCTGCAGACAGTCTCCATCAATGACAGGCACTCTGGCGCTACAGC 6682
Db GATGGGACGGGAGCTCCAGAGCGTGGCGCTCAATGACCGCCGACCTGGCGCTACAGC 6686
QY 6683 TAGACCTCAATGGGAACCTTGCACTTACTGAGCCCTCGGAACAGTGCACCGCTCACCA 6742
Db TATGACCTTATGGGAATCTCCACTTACTGAAACCCAGGCAACAGTGTGCGCTCATGCCC 6746
QY 6743 CTACGGTATGACATCCCGGACCGCATCACTCGGCTGGGTGACGTGCAATACAGATGGAT 6802
Db TTGCGCTATGACCTCCGGATCGGATACAGACCTCGGGATGTCAGTACAAATGAC 6806
QY 6803 GAGATGGCTTCTGAGCAGCGGGCGGTGATCTTTGAGTACAACTCAGCTGGCGTG 6862
Db GACGATGGCTATCTGTGCCAGAGAGGTCTGACATCTTCGAATACAACTTCAAAGGCGCTC 6866
QY 6863 CTGATCAGGCTACACCGGCTGGCAGCTGGAGTGTCAAGTACCCTGACGTAGGCTG 6922
Db CTAAAGAGGCTTAAACAAGGCGAGCGGTGGAGTGTCCAGTACCGCTATGATGGCGTA 6926
QY 6923 GGGCGGCGGTCTCCAGCAAGAGCAGCACAGCCACCACTGCAAGTCTTCTATGACAGAC 6982
Db GGAAGCGGGCTTCTACAGACCAACCTGGGCCACCACTGCACTTCTACTCTGAC 6986
QY 6983 CTGACCAACCCCAACAAGGTCAACCACTGTACAACTCCAGCTCTGAGATCACTCC 7042
Db CTCACAAACCGGAGCGCATCACCATGTCTTACAACTCACTCCAACTCGGAGATTAACCTCA 7046
QY 7043 CTCTACTACGATTTGCAAGGACACTCTTTGCCATGAGCTGAGCAGTGTGATGATTT 7102
Db CTGTACTACGACCTCCAGGGCCACTCTTTGCCATGAGAGCAGCAGTGGGAGGATAC 7106
QY 7103 TACATAGCTTGTGACAACTCGGACCCCTCTTCTGTCTTTTGTAGTGAACAGGTTTGATG 7162
Db TATGTTGCTCTGTATAACAGGAGCTCTCTGCGTGTGTTTCACTCAACAGGCTCATG 7166
QY 7163 ATCAGCAATCTGTACAGCCTATGGGAGATCTACATGATACCAACCCCACTTT 7222
Db ATCAACAGCTGCAGTACAGGCTATGGGAGATTTATTATGACTCCAAACCCCGACTTC 7226
QY 7223 CAGATCATATAGGCTACCATGTGGCTCTATGATCCACTCACCAGCTGTGCCACATG 7282
Db CAGATGCTATGCTTCCATGGGGACTCTATGACCCCTGACCAAGCTGTCTCACTTC 7286
QY 7283 GGGCGGCGAGTATATGATGCTGGCGGAGCTGTGACTAGCCCGACACCGAGCTGTGG 7342
Db ACTCAGCGTATATGATGCTGGCAGGACGATGGACCTCCCGAGACTATACCATGTGG 7346
QY 7343 AAGCACCTTATGACAGCAACGTCATGCTTTTAACTCTATATGTTCAAAACAACAAC 7402
Db AAAAAGTGGCAAGGAGCGCGCC---CCCTTTAACTGTATATGTTCAAGGACAACAT 7403
QY 7403 CCCATCAGCAATCCCAGGACATCAAGTGTCTTATGACAGATGTTAACTGCTGCTC 7462
Db CCTCTCAGCAGTGTAGTATTTGAAGAACTACGTGACAGATGTGAAAAGCTGGCTGTG 7463
QY 7463 ACCTTTGGATTCAGCTACAAAGTGTACCTGGTTATCCCAACACAGATGATGCC 7522

Db 7464 ATGTTTGGATTTTCACTTTAGCAACATCATCTCTGCTTCCCGAGAGCCAAATATGATTTTC 7523
QY 7523 ATGGAACCTCTCTAGAGCTCATCACACAGATGAATAACGAGAGTGGGACACAGC 7582
Db GTGCTCTCTCCCTATGA-----ATTGTGAGAGTCAAGCAGTGAATGGA 7571
QY 7583 AAGTCTATCTCGGGTACAGTGTGAAGTACAGAGCAGCTCAAGCCTTTGTTCACCTTA 7642
Db CAGCTCATTTACAGTGTCCACACAGACAACAGAGAGACATAACAGGCTTTCATGGCTCTG 7631
QY 7643 GAACGGTTTGACAGCTCTATGGCTCCCAATCACAGCTGCCAGAGCTCCAAAGACC 7702
Db GAA-----GGACAGGTCTATTACTAAAGCTCCAGCCAGCATCCGAGAGAAACAGGT 7685
QY 7703 AAGAAGTTTGTGATCCAGCGGCTCAGTCTTTGGCAAGGGGGTCAAGTTTGCCTTTGAAGAT 7762
Db CACTGGTTTGCCACCAACAGCGCTCATTTGGCAAGGATCATGTTTGGCCATCAAGAA 7745
QY 7763 GCGCAGTACACACAGATCATCATGCTGGCCATAGAGATGGGCGAGGGTTCCTGCC 7822
Db GGGCGGGTGAACACAGCGGGCTGTCCAGCATCGCAAGAGATAGCCGCAAGGTGTCATCT 7805
QY 7823 ATCTGAACCATGCGCACTTACCTAGAGAACTGTCACTTCACTTGTATGGGTGATATACC 7882
Db GTGCTGAACAGCGCTTACTCTCTGACAAAGATGCACTACAGCATCGAGGCGAAGACACC 7865
QY 7883 CATTACTTTGTGAACACAGGACCTTTCAAGAGTGAACCTGGCCATCTCTGGGCTCTAGTGG 7942
Db CACTACTTTGTGAAGATTGGCTCAGCGATGGCGACTGGTCACTACTAGCCACCCATC 7925
QY 7943 GGGCGCGGAACCTCGAGATGGGGTCAACGTCACCTGTCTCCAGATCAACACAGTACT 8002
Db GCGCGAAGTCTAGAGCGGGTGAACGTGACCGTGTCCAGCCCAACCGTGTGTC 7985
QY 8003 AATGCGAGGACTAGACGCTTACACAGACATCCAGCTCCAGTACGGGGCACTGTGCTTGAAC 8062
Db AACGGCAGGACTCGAAGGTTTCAAGAACTTGAAGTTCCAGTACTCCACGCTGTGCTCAGC 8045
QY 8063 ACACGCTACGG-----GACAAGCTTGGATGAGAGAGAGGACGCGGTCTCTGAGCTG 8113
Db ATCCGCTATGGCTCACCCCGACACCTTGGACGAAGAGAGGCGCGGTCTCTGACACAG 8105
QY 8114 GCGCGCAGAGAGCGCTGCGCAAGCTGGGCGCCGAGCAGCAGACAGACTGCGGGAAGGG 8173
Db GCGAGACAGAGGCGCTTGGGCGCGCTGGGCGAGGAGCAGAGCAAGCCAGGACGGG 8165
QY 8174 GAGGAAGGCTCGGGCTTGGACAGAGGGGAGAGAGCAGAGTGTCTGAGCAGAGGCGG 8233
Db AGAGAGGGGAGCGGCTGTGAGCTGAGGCGGAGAGCAGAGCTTCTGAGCACCGGGCGC 8225
QY 8234 GTGCAAGCTACAGCGCTTTTTCGTGATCTCTGTCAGAGCTACCCAGAACTGTCAAGAC 8293
Db GTGCAAGGTTACAGGGATTTTACGTCTTCCGTTGGAGCAATACCCAGAGCTTGCAGAC 8285
QY 8294 AGCGGCAACAACATCCATCTATGACAGAGCAGATGGGCGGAGGTGACAGA 8348
Db AGTAGCAGCAACATCCAGTTTAAAGACAGATGAGATGGGAAGAGGTAAACAAA 8340

RESULT 15

ACC72051
ID ACC72051 standard; DNA; 9058 BP.

XX

AC ACC72051;

XX AC

DT 08-JUL-2003 (first entry)

XX

DE BCU0205A gene #SEQ ID 79.

XX

KW Breast cancer; cytostatic; gene therapy; antisense therapy; regulated;

KW drug discovery; clinical medicine; forensic medicine; gene;

KW chromosome 5q33.3; ds.

XX Homo sapiens.
OS WO2003029421-A2.
FN XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031287.
XX
PR 03-OCT-2001; 2001US-0326536P.
PR 14-MAY-2002; 2002US-00144194.
XX
XX (ORIG-) ORIGENE TECHNOLOGIES INC.
XX
XX Sun Z, Li X, Fan W, Kovacs KF, Jay G;
XX
XX WPI; 2003-381623/36.
XX
XX P-PSDB; ABR58317.
XX
XX New isolated human differentially-regulated breast cancer polynucleotide
XX and polypeptide, useful for diagnosing, staging, prognosticating,
XX preventing and/or treating diseases and conditions relating to breast
XX cancer.
XX
XX Claim 2; SEQ ID NO 79; 127pp + Sequence Listing; English.
XX
XX The invention relates to isolated polynucleotides which are
XX differentially-regulated in breast cancer. The methods and compositions
XX of the present invention are useful for detecting, diagnosing, staging,
XX monitoring, prognosticating, preventing and/or treating diseases and
XX conditions relating to breast cancer, and may be used in gene therapy or
XX antisense therapy. They can also be used in research, drug discovery,
XX clinical medicine and forensic medicine. Sequences given in records,
XX ACC72012-ACCT2074 represent polynucleotides of the invention that are
XX differentially-regulated in breast cancer. NOTE: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 9058 BP; 2387 A; 2430 C; 2263 G; 1978 T; 0 U; 0 Other;
XX
Query Match 36.7%; Score 3068.6; DB 7; Length 9058;
Best Local Similarity 65.3%; Pred. No. 0;
Matches 4890; Conservative 0; Mismatches 2264; Indels 333; Gaps 13;
XX
880 CGGGCACTTCTCTTCAAG--CTGAGGACCTCCCGGCTCTTCTGCACCATCAC 936
DB 532 CCGGCACTTCTCTTCAAGACCTCTCGGGAGGACACACCTTGTTCAGAGCTTCC 991
XX 937 AGGTATCCACTGACGTCCAGACAGTGTACTCTCTCCGCCCGGACCCCTGCCG 996
DB 592 GGGATACCTTTGACCTCAGAAACGGTTTACAGCCCGCCCGGCTGTGCCAGGAA 651
XX 997 CACCTTCCCGCCCGGCTTACCTCAGAGAGCCCTCCAGTACTGTACTGGAAGTG 1056
DB 652 TACTTTCTCAGGAGGCTTCAAGCTGAAGAGCCCTCCAAATCTGACGCTGGAATG 711
XX 1057 CGCAGCCCTGAGCCGCAATCTCATCTCAGCCACTCTGGTTCATCTGTGGCATPACTTGT 1116
DB 712 TGCTGCCCTCTCCGCCAATCGCGGCGCTCTCTTGGCTATTTTGTGCGTATTTCAT 771
XX 1117 GGCATGCACTGTTGGCTTAACCTGACCTGACCGCATGAGAGGAGGAGATGATGA 1176
DB 772 AGCAATGCACTGCTCGGATCTCAATTTGGCACTCCAGCTGACAGTGGGACACCTTAA 831
XX 1177 GATCAGGAGGACACAGCCAGAGTGGCCCTGTGCCAACCGAGCTCTCCCTATACCCCTC 1236
DB 832 CA-----ATGGGATAAGGACCGGCTTACAGGAAACGATGATGGCAACATGCCATC 885
XX 1237 AGGGGGCACTGGCTTGAAGACCCCTGACAGGAAAGGCAAGGAACACACAGAGAAAGCC 1296
DB 886 TGGAGGCAAAAGTGCCCTGGTGGTTGAA----- 914

QY 1297 CAGTAGTTTCTTTTCCAGAGGACAGTTTTCATAGATCTTGGAGAAATATGATGTGGGAAGCG 1356
DB 915 -----ACAGCAGCATAGACAGTGGTGAAGCAGAAGTTGTGCGCG 954
XX
QY 1357 AGCCTCCAGAAAGATTCTCTGGCACATTCTTGAGATCTCAAGTGTTCATAGACCATCC 1416
DB 955 GGTAAACACAGAGTCCACACAGGGGTGTTTGGAGGTCAAAATTCACATCATCAGCC 1014
XX
QY 1417 TGTGCATCTGAAATTCATATGTCTCTGGGAAAGGACGCTCTGTGTGGCATTTAGCGAG 1476
DB 1015 CCAGTTCTTAAAGTTCACATCTCCCTCGGAAGGAGCTCTCTTGTGTGTTTACATAAG 1074
XX
QY 1477 AAAAGGCTCTCTCTTCCATACATACAGTTTGTGAGTCTTGTGGAGCTGTGGATGGCAGG 1536
DB 1075 AAGAGGACTTCCACATCTCATGCCAGTATGACTTTCATGGAAGCTCTGGACGGGA---- 1130
XX
QY 1537 GCTCTAAACCCAGGAGGCGGAGCCTTAGAGGGACCCCGCGCCAGTCTCGGGGAATGT 1596
DB 1131 -----AGGAGAAAGTGGAGTGTGTTGAGTCTCCAGGGAACGCCGAGCATACA 1179
XX
QY 1597 GCCCCCTCCAGCCATGAGACAGGCTTCACTCAGTATTTGGAATTCAGGAATCTGGCACTT 1656
DB 1180 GACCTTGGTTTCAAGATGAAGCCGTGTTGTGACGATCTGAGATGGGCTGTGGCACT 1239
XX
QY 1657 GGCCTTTTACATGACGGAAGGAGTCAAGAGTGGTTTCTCTTCTCACCACCTGCATTA 1716
DB 1240 GGCCTTCTACATGATGGAAGACAAAGAGATGGTTTCTTCAATCTGTTGCTCTAGA 1299
XX
QY 1717 GTCGGTGAATAACTCCCGACCACTGCTATGCAATGGTACTGCACTCTCTGGGACCTG 1776
DB 1300 TTCAGTGCAGGACTGTCCACGTAACCTGCAATGGGAATGGTGAATGTGTCCGGGGTGTG 1359
XX
QY 1777 COACTGCTTCTGGTGTTCCTGGGCCCGGCTGTGGGAGAGCCCTCTCCCGGCTCTG 1836
DB 1360 TCAGTGTCTCCAGATTTCTAGGAGCAGACTGTGCTAAAGCTGCCCTGCTCTGCTGTG 1419
XX
QY 1837 TAGCGGAAATGGCAATACATGAAAGGCAGATGTTGTGCCACAGTGGCTGGAAAGCGC 1896
DB 1420 CAGTGGGAATGGCAATATTTCTAAAGGACGTCGACGCTGTACAGCGCTGGAAGGTGC 1479
XX
QY 1897 TGAGTGCATGTGCCACCAACACAGTGTATGATGTGGCTGTGACAGCAACCATGCACTG 1956
DB 1480 AGAGTGCAGTGTGCCATGAAATCAGTGCATGATCTTCTTCTCGGGGGGCGCACGCTCTG 1539
XX
QY 1957 CATCACGGGCACTGTGATCTGCAACCTTGGCTACAAGGGGAGAGCTGTGAGGAAGTGA 2016
DB 1540 CATTCATGGGAACGTGTCTGCTGTGCTGTGCTACAAAGGGGAGGACCTGTGAGGAATGA 1599
XX
QY 2017 CTGATGACCCCAACATGTTTCCAGCCCGGGGTGTGCTGTGAGAGGCGCAATGCTTGCCT 2076
DB 1600 TTGCTTGGATCCCACTGCTCCAGCCAGGAGTGTGTGTAATGAGAAATGCTCTGTGCGAG 1659
XX
QY 2077 TGTGGATGGGAGGACCACTCGAGAGACCCCGGAGGCGGAGAGCTGTGAGGAAGTGTTC 2136
DB 1660 CCCTGGCTGGGTGTGTGAACTGTGAGTGGCGAGGCTCAGTGTCCAGACCAAGTGCAG 1719
XX
QY 2137 AGGCCACGGAACCTTCTCTCCCGGACACCGGGCTTTGAGCTGTGACCCAGCTGGAATGG 2196
DB 1720 TGGGCATGGCAGTACCTGCTGACAGCGGCTCTGACAGCTGCGATCCCAACTGGATGGG 1779
XX
QY 2197 ACAGACTGTTCTCAGATCTGTGCTGCCGACTGTGTTGGCCATGGCTGTGCTGAGG 2256
DB 1780 TCCGACTGCTCTGTTGAT----- 1798
XX
QY 2257 GGGCACCTTCCGCTCGAGGAGTGGCTGGATGGGGGAGCCCTGCGACACGCGGCGCTGCCA 2316
DB 1799 ----- 1798
XX
QY 2317 CCCGGCTGTGCCGAGCATGGACCTCCCGGAGCGCAAGTGGGAGTGCAGCTGCGCTG 2376
DB 1799 ----- 1798
XX
QY 2377 GAATGGGGAACACTGACCATCGCTCTACTATCTGATAGGTTGTTAAAGAGGGTTGCC 2436

Db	1799	-----GGTGGCC 1806
Qy	2437	TGGTGTGCAATGGGCAACGGCAGATGTACTTTAGACCTGAATGGTTGGACCTGCGTCTG 2496
Db	1807	TGACTTTGCAACGGTAAACGGAGATGACACACTGGGTGAGAGAGTGGCAGTGTGCTG 1866
Qy	2497	CCAGCTGGGTGGAGAGAGCTGGCTGTGACACTTCCATGGAGACTGCCCTGGCGTGACAG 2556
Db	1867	CCAGACGGCTGGAGAGGGCCGGATGCAAGTTGCCATGGAACTTCTGTGCTGATAA 1926
Qy	2557	CAAGAACAATGATGGAGATGGCTGGTGGACTGTGATGGAACCTGACTGTGCTCCAGCC 2616
Db	1927	CAAGTAATCAGGAGATGGCTGGTGGATTTGTTGGAACCTGACTGTGCTGGCTGAGTC 1986
Qy	2617	CCTGTGCCATATCAACCGCTGTGCTTGGCTCCCTTCCCTTACCTTGGACATCATCCAGGA 2676
Db	1987	AGCCTGTGACAGACCTGTCTTCCCGGGGTCCCGGACCACTGGACATCAITCAG-- 2044
Qy	2677	GACACAGTCCCTGTGTACAGCAGAACTTACACTCCTTCTATGACCGCATCAAGTTCT 2736
Db	2045	----CAGGGCCAGACGATTGGCCCGCAGTGAAGTCTTCTATCAGCGTATCAAGCTCTT 2100
Qy	2737	CGTGGGACGGACAGCAGCACATAATCCCGGGGAGAACCCCTTTGATGGAGGATGC 2796
Db	2101	GGCAGGCAAGATAGACCCACATCATTTCTGGAGAGAACCTTTCAACAGCAGCTTGGT 2160
Qy	2797	TTGTGTTATTGCTGGCCAAAGTGATGACATCAGATGGAAACCCCTGGTTGGTGTGAACAT 2856
Db	2161	TTCTCTCATCGAGGCCAAGTAGTAATACTACAGATGGAACCTCCCTGGTGGTGAACGT 2220
Qy	2857	CAGTTTGTCAATAACCTCTCTTTGGATATACATCAGCAGGCAAGATGGCAGCTTTGA 2916
Db	2221	GTCTTTTGTCAAGTACCAAAATACGGCTACACCATCACCGCCAGGATGGCAGTTGGA 2280
Qy	2917	CTTGGTGACAAATGGCGCATCTCCATCTCGGTTTCGAGGGGACCTTTTCATCAC 2976
Db	2281	CCTGATCGAAATGAGGTGCTTCTTGACTCTACATTTGAGCGAGCCCGTTTCATGAG 2340
Qy	2977	ACAGGACACACCTGTGGTGCATGGATCGCTTCTTTGTCAATGAAACCAATCATCAT 3036
Db	2341	CCAGGAGCGCACTGTGTGGCTGGCTGGACAGCTTTTACGCCATGGACACCCCTGGTGT 2400
Qy	3037	GAGACATGAGAGATGAGATCCAGCTGTGACTGACAGCAATTTTCCCGCCCAACCC 3096
Db	2401	GAAACCGAGAGAACTCCATCCCGAGCTGTGACTCAGTGGCTTTTTCGCGGCTGATCC 2460
Qy	3097	AGTGGTCTCTCCATCCCACTGAGTGCCTTCCCGCAGCTCTGTGACAGAGAAAGGCCCAT 3156
Db	2461	AATCATCATCTCTCCCACTGTCCACCTTCTTTAGTGTGCCCTGGCAGAAATCCCAT 2520
Qy	3157	TGTGGCGGAAATCAGGCTTTGAGAGGAAATCTCTATCTGGCTGGAGATGAGGCT 3216
Db	2521	CGTGGCTGAGACCCAGGTTCTTCATGAAGAAATCGAGTCCCTGGTTCCAATGTGAACT 2580
Qy	3217	GAGCTACTGAGCAGCGGACCCCTGGCTACAAATCTGCTGAGGATCAGCTCACCCA 3276
Db	2591	TCGCTATCTAGCTCTAGAACTGAGGGTACAGGTACAGTCACTGCTGAAGATCACCATGACCCA 2640
Qy	3277	CCGACCACTCCCTTCACTCAAGGTCACCTCATGATAGCGTGGAGGGCCGCT 3336
Db	2641	GTCCACAGTGGCCCTGAACCTCAITAGGGTTCACTGATGGTGGCTGTGAGGGGCATCT 2700
Qy	3337	CTTCAGGAGTGGTTCGCTGAGCCCGACAGCTGTCTTATTTTCAATTTGGGACAGAC 3396
Db	2701	CTTCCAGAAATCATTCAGGGCTTCTCCCACTGGCTTACACCTTCATCTGGGCAAGAC 2760
Qy	3397	AGAGCTCTACAAACAGAGGTGTTTGGGCTTTTCAGAGCCCTTTGTTCCGTTGGTTATGA 3456
Db	2761	AGATCGTATGGCAAGGGGTATAGGACTCTCAGATGCTGTGTGTCTGTGCGGTTTGA 2820
Qy	3457	ATATGAATCTGCGCCAGATCTAATCCTGTGGGAAAGAACACAGCTGTGAGGGCTA 3516

Db	2821	ATATGAGACCTGTCCCACTTAATTTCTGGGAGAAAGGACAGCCCTCTTTCAGGGATT 2880
Qy	3517	TGAATTCAGCGCTCCCAAGCTTGGAGGATGAGCCTAGACAAACATCATGCCCTCAACAT 3576
Db	2881	CGAGCTGAGCCCTCCCACTCGGTGGTGGTCTCTAGAGAAACACACATCTCTCAATGT 2940
Qy	3577	TCAAAGTGGTATCTCTGCAAAAGGAATGGGGAGAACAGTTTGTGTCTAGCAGCCTCC 3636
Db	2941	TAAAAGTGAATCTCTACAAAGGCACTGGGGAAACCAAGTTCTCTGACCCAGCAGCCTGC 3000
Qy	3637	TGTCAATGGGAGCATCATGGGCAATGGGCGCGGAGAGCATCTCTCCGCCAGCTGCAA 3696
Db	3001	CATCATACCAGCATCATGGGCAATGGTCCGCCCGGAGCATTTCTCTCCAGCTGCAA 3060
Qy	3697	CGGCCCTTGTGACGGCAACAAAGCTCTCGCCCCAGTGGCCCTCACCTGTGCTCTGACGG 3756
Db	3061	CGGCCCTTGTGAAGCAACAAAGCTCTCGCCCCAGTGGCTCTGGCTGTGGAATCGATGG 3120
Qy	3757	GAGCCTCTATGTGGTGAATTTCAACTACATTAGAAGGATCTTCCCTCTGGAATGTAC 3816
Db	3121	GAGCCTCTATGTGGTGAATTTCAATTACATCGAGCATCTTCCCTCTCGAAATGTGAC 3180
Qy	3817	CAACATCTTAGAGCTGAGGAATAAAGATTTTCAAGATAGTCACAGTCCAGCACACAAATA 3876
Db	3181	CAGCATCTTGAATTTACGAATTAAGATTTAAACATAGCAACACCCAGCACACAGTA 3240
Qy	3877	CTACTTGGCCACAGACCCCATAGTGGGGCGCTTTCTCTTTCTGACAGCAACAGCGCG 3936
Db	3241	CTACTTGGCAGTGGACCCCGTCTCGGCTCGCTTACGTTCGAGCACCAACAGCAGGAG 3300
Qy	3937	GGTCTTTAAATCAAGTCCACTGTGTGGTGAAGGACCTTGTCAAGAACTCTGAGTGT 3996
Db	3301	AACTACCGCTCAAGTCTCTGAGTGGAAACCAAGACCTTGGCTGGGAATTCGGAATGT 3360
Qy	3997	TGCGGGGACAGGTGACCACTGCTCCCTTTGATGACACTCGCTCGGGGATGGTGGGAA 4056
Db	3361	GGCAGGAGCGGAGAGCAGTGTCTACCTTTGATGAAGCCCGCTCGGGGATGGAGGAA 3420
Qy	4057	GGCCACAGAGCCCACTCACCAATCCAGGGGTATTACAGTGGAGCAAGTTTGGCTGAT 4116
Db	3421	GGCCATAGATGAACCCCTGATGAGCCCGAGAGTATTCAGTAGAACAAAGATGGGCTCAT 3480
Qy	4117	CTACTTCTGGTGGACCACTCATGATCAGACGCACTCGATCAGAAATGGGATCATCTCCACCT 4176
Db	3481	GTACTTTGTCTGATGCCACCATGATCGGAAGTTGACCAAGATGGAATCATCTCCACCT 3540
Qy	4177	GCTGGCTCTAATGATCTCATACAGCCCGGCACTCAGCTGTGATTCTGTGATGATAT 4236
Db	3541	GCTGGCTCTCAATGACTCACTCGCTCGGCGCTGAGCTGTGATTCCAGCATGGATGT 3600
Qy	4237	TTCCAGGTAAGACTGGAGTGGCCACAGACTTAGCCATCAACCAATGGCAACTCACT 4296
Db	3601	AGCCAGGTTCTGTGGAGTGGCCACAGACCTTGTGTCAATCCATGGATCACTCTT 3660
Qy	4297	TTATGTCTCGAACAATGTGTCTCTGCAAACTCTGAAACACCAAGGTGGCAATTTT 4356
Db	3661	GTATGTTCTAGAGAACAAATGTCTCTGAAATCAACGAGAACCAAGTCAAGTCAATCAT 3720
Qy	4357	CGCCGGAGGCCATGCACTGCCAGTCCCTGGCATTGACACTTCTCTGAGCAAGT 4416
Db	3721	TGCGGACGCCCATGCACTGCCAAGTCTCTGGCATTTGACTACT---CACTCAGCAACT 3777
Qy	4417	GGCCATCCAGCAACCCCTGGAGTCAGCCACCGCTTGGCTGTTTTCACAAATGGGGTCTCT 4476
Db	3778	AGCCATCTACTCTGCCCTGGAGTCAGCAGTGCCATGGCATTTCTCACACTGGGGTCTCT 3837
Qy	4477	GTATATGTGAGATGAGAAAGATCAACCGCATCAGGAGGTCAACACTAGTGG 4536
Db	3838	CTACATCACTGACAGATGAGAAGATTAACCGTCTACGCCAGGTAAACCAACCG 3897
Qy	4537	AGAGATCTCAGTGTGGGGCCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTG 4596
Db	3898	GGAGATCTGCTTTTAGCTGGGGCAGCTCGGACTCGGACTGCAAAAAAGATGTCAATTTG 3957

QY 4597 TGATGTTTTCTGGAGACGATGTTATGCCAAGATGCAAAAGTTAAATACCCCATCTTC 4656
 Db 3958 CAATGCTATTTCAGAGATGATGCTACGGACATGATGCCATCTTGAAATTCCTCCATCATC 4017
 QY 4657 CTTGGCTGTGTGCTGATGCGGAGCTCTACCTGGCCGACCTTGGGAACATCCGAATTCG 4716
 Db 4018 CTTAGCTGTAGCTCCAGATGATCAATTTACATTCAGACCTTGGAAATATTGGATCAG 4077
 QY 4717 GTTTATCCGGAAGAACAGACCTTTCTCAACACCCAGAACATGATGATGCTTCAACC 4776
 Db 4078 GCGGCTCAGCAAGAACAGACCTTTCTTAATGCTTCAAGCAGTATGAGGCTGCATCCCC 4137
 QY 4777 AATTGACAGGAGCTCTATCTGTTTGATACACCGGCAAGCACCCTGTACACCCAAAGCCT 4836
 Db 4138 CGGAGAGCAGAGTTATATGTTTTCAAGCTGATGCGATCCACCAATACACTGTGAGCCT 4197
 QY 4837 GCCACAGGAGACTACCTGTAACTTCACTACACTGGGAGCGGACATCACACTCAT 4896
 Db 4198 GGTGACAGGGAGTACTTGTACAATTTTACATATAGTACTGACAAATGATGCTCACTGAAT 4257
 QY 4897 CACAGCAACATGGCAACATGTTAAATGTCGCGGAGACTCTACTGGGATGCCCCCTCG 4956
 Db 4258 GATTGACAATTAATGGGAATTCCTCGAAGATCCGTCGGGACAGAGTGGCATGCCCGTCA 4317
 QY 4957 GCTGGTGTCCAGATGGCCAGGTGTACTGGGTGACCATGGGACCAACAGTGCACCTCAA 5016
 Db 4318 CTTGCTATGCTGACAAACAGATCATCACCTCACCGTGGGACCAATGAGGCGCTCAA 4377
 QY 5017 GAGTGTACACACAGGACACAGATGTTGGCCATGATGACATACCATGGGCAATTCGGGCT 5076
 Db 4378 AGTGGTGTCCACACAGAACCTGGAGCTTGTCTCATGACCTATGATGGCAACACTGGGCT 4437
 QY 5077 TCTGGCAACAAAGCAATGAAACCGATGGACAAACATTTTATGATGACGACAGCTTTGG 5136
 Db 4438 CTTGGCCACCAAGAGCATGAAACAGGATGGACGATTTCTATGATGATGACCAAGG 4497
 QY 5137 CCGCTGACAAATGTGACCTTCCCTACTGCGCAGGTGAGCAGTTTCCGAAGTATACAGA 5196
 Db 4498 CCGCTGACCAACGTGACGCGGCCCCACCGGGGTGGTAACAGTCTGCAACCGGAATGGA 4557
 QY 5197 CAGTTGATGATGTCAGATGAGAGCTCCA---GCAAGATGATGTCACCATACAC 5253
 Db 4558 GAAATCTATTACCATGACATGAGAACTCCAACCGGATGATGACGTCACTGTCTATCAC 4617
 QY 5254 CAACCTGTCTGCTCAGGCGCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTA 5313
 Db 4618 CAACCTGTCTCAGTAGAGGCTCTTACACAGTGGTACAAGATCAAGTTCCGGAACAGCTA 4677
 QY 5314 CTACATCGGGCGCATGGCTCTTGGGCTGCTGCTGGCCAAACGGATGAGGTGGCGCT 5373
 Db 4678 CCACTCTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4737
 QY 5374 GCAGACTGAGCGCCCACTGCTGGTGGCACCCTGCAACCCCGTGGGCAAGGAAATCT 5433
 Db 4738 CCACAGGAGGCCCATGCTCTAGGGGACCATCAACCCCATTTGGAGCTGCAACAT 4797
 QY 5434 CAGCTGCCATGCAACAGGCTTCACTGGTGGAGTGGCGCCAGCGCAAAAGAGCAGCG 5493
 Db 4798 CTCCTGCTATGAGAAATGGCTTAACTCCATTGATGGGCGCTTAAGAAAGGAAACAGAT 4857
 QY 5494 TCGGGGCGAGCTCACTGCTCTTGGGCGCGCTGCGGTGGCAACCCGAATCTCTCTATC 5553
 Db 4858 TAAAGGCAAGTCAACATCTTGGGAGAGCTTCCGGGTCATGGAAGAAATCTCTTGTG 4917
 QY 5554 TCTGGACTTGTATGCGGTAAACGCAAGAGAGATCTATGATGACCAACCGCAAGTTTCA 5613
 Db 4918 CAATTGATGATGCAAAATTCGGACTGAAAGATCTATGATGACCAACCGCAAGTTTCA 4977
 QY 5614 CTTTCGATTTCTGACACAGGCGGGCGGCGGCGGCTGCTGTCACCCAGCAGAGGCT 5673
 Db 4978 CTTGAGGATCAATTTATGACAGGTGGCGCGGCGGCTTCTCTGCTGCCAGCAGCGGCT 5037

QY 5674 GAATGGTGTCAACGTGACATCTCCCTTGGGGTTTACATTTGCTGGCATCCAGAGGGCAT 5733
 Db 5038 GGCAGCTGTCAACGTGTATATCTTCAATGGGCGCTGGCTGGCTTCAGCGTGGGCG 5097
 QY 5734 CATGTCTGAAAGATGGAATACGACAGCGGGCGGATCACATCCAGGATCTTCGCTGA 5793
 Db 5098 CATGAGCAGAGGACAGACATCGCAACAGGCGCATCGTCCCGCATGTTCCGCTGA 5157
 QY 5794 TGGGAAGACATGAGAGCTACATATCTTGAAGAGTCCATGCTGTCTTACTACAGCCA 5853
 Db 5158 CGGGAAGTGTGGAGCTACTCTACTTGAAGTCCATGCTCTCTCTGCTTCAGAGCCA 5217
 QY 5854 GAGGAGATATCTTGTAGTTGCGAAGAAAGACCGCTCTTCTGTGAGAGTCCCAA 5913
 Db 5218 AGTCAAGTATATTTGATGATGATCTCTCTGACCGCTCTTGGCTCAGTACCAAGCCAG 5277
 QY 5914 CTTGGCGGCGAGACATPAGAGACCTCCGCTCAGTGGGCTTACTACAGAAACATCTATCA 5973
 Db 5278 CTTGGCGGCGAGACATGTCACACACACCTCCATGGCTTACATCCGTAATATTTACAA 5337
 QY 5974 GCCCTGAGGCGAATGCTCAGTATACAGAGCTTCTACTGAGATGGGACCTCTCTCA 6033
 Db 5338 CCGCTTGAAGCAATGCTTCGGTCACTTTGACTACAGTATGACCGCGCATCTCTGAA 5397
 QY 6034 CACCTTCTACCTGGGCACTGGCGGAGGTGATATCAAGTATGGCAAACTGTCAAAGCT 6093
 Db 5398 GACCTCTTTTGGGACCGGACCGGAGTGTCTTCAAGTATGGAAACTCTCAAAGT 5457
 QY 6094 GCGAGAGAGCTCTATGACACCAAGGTGAGTTCACCTATGACAGAGCGGAGGAT 6153
 Db 5458 ATCAGAGATTTCTACGACAGTACCGCGCTCACCTTGGGTATGACGAGACCACTGGTGT 5517
 QY 6154 GCTGAAGACCACTCAACTACAGATGAGGCTTCCCTGACCATCCGCTACCGTCAAGT 6213
 Db 5518 CTTGAGATGTCACCTCCAAAGTGGGCTTCTCTGACCAATGAGTACCGAAGAT 5577
 QY 6214 TGGGCGCTGATGACGACAGATCTTCCGCTTCACTGAGGAAGGATGTCACCGCG 6273
 Db 5578 TGGCGCTGTTGGACAAAGAGATCTACAGTCTCCGAGGAAGGATGTCATGCGCAG 5637
 QY 6274 TTTTGAATCAACT---ATGACACAGCTTCCGGTGACGAGCATGAGGCTGTGATCA 6330
 Db 5638 GTTTGACTACACCTATCATGACAAACAGCTTCCGATCGAAGCATCAAGCGCGTCAAG 5697
 QY 6331 CGAGACCCACTGCGCAATGATCTCTATGCTATGATGATGTCAGGCAAGACAGAGCA 6390
 Db 5698 TGAGACTCCCTCCCGTTGACCTTACCGCTATGATGATGATTTCTGGCAAGGTGGAACA 5757
 QY 6391 GTTTGGGAGTTTGGTGTCTTACTATGACATTAACAGATCATCACCAGCTGTCTAT 6450
 Db 5758 CTTTGGTAAAGTTTGGAGTCACTATTATGATCAATCAACAGATCATCACCAGCTGTAT 5817
 QY 6451 GACCCACACCAAGCATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 6510
 Db 5818 GACCTCAGCAAAACACTTCGACACCCATGGGCGATCAAGAGGTCCAGTATGAGATGTT 5877
 QY 6511 CCGTCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6570
 Db 5878 CCGTCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5937
 QY 6571 GGAGCTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6630
 Db 5938 GGAGCTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5997
 QY 6631 CCGGCGCTGACAGAGCTTCCATCAATCAAGCAAGCTTCTGGGCTTACAGCTACGACCT 6690
 Db 5998 CCGGCGCTGACAGAGCTTCCATCAATCAAGCAAGCTTCTGGGCTTACAGCTTACCT 6057
 QY 6691 CAATGGGAACCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 6750
 Db 6058 TAATGGGAATCTCACTTACTGAAACCGGCAACAGTGTGGCTCATGCGCTTGGCTA 6117
 QY 6751 TGACATCCGCGACCGCATCTACTCGGTGGGTGAGTGCATTAACAAGATGATGAGGATG 6810

[illegible]

Db	7177	CAACGCCTACTACTGGACAAGATGCATCAGCATCGAGGCGAAGACACCCACTACTT	7236
Qy	7891	TGTGAACACGAGACCTTCAGAAAGTGAACCTGGCCATCTCGTGGGGGGCGCG	7950
Db	7237	TGTGAAGATTGGCTCAGCCGATGGCGACCTGGTCACACTAGGCACCACTATCGGCCGCAA	7296
Qy	7951	AACCCGTGGAGAATGGGTCAACGTCACATGTGTCACAGATCAACACAGTACTTAATGGCAG	8010
Db	7297	GGTGCTTAGAGAGCGGGGTGAACGTGACCGTGTCCAGCCCAAGCTGTGTCTCAACGGCAG	7356
Qy	8011	GACTAGACGCTACACAGACATCCAGTCCAGTAGGGGCATGTCTTGAAACACACGCTA	8070
Db	7357	GACTCGAAGTTTCAGAAACATTGAGTTCAGTACTCCACGCTGTCTCAGCATCCGCTA	7416
Qy	8071	CGS-----GACAACTTTGGATGAGAGAGGACAGCGGTCTCTGGAGCTGGCCCGGCA	8121
Db	7417	TGCGCTTCAACCCCGCACACCTCGACAGAGAGAGCCCGCGTCTCTGGACCAAGGCGAGACA	7476
Qy	8122	GAGAGCGCTGCGCAAGCGTGCGGCCCGGAGCAGCAGAGACTGCGGGAAGGGGAGGAGG	8181
Db	7477	GAGGGCTCTGGGCACGGCCCTGGCCCAAGGAGCAGCAGAAAGCCAGGACGGGAGAGAGG	7536
Qy	8182	CTTGCGGGCCTGGAACAGAGGGGGAGAACGACGAGTGTGTAGCAACAGGCGGTTGCAAGG	8241
Db	7537	GAGCCGCGCTGTGAGCTGAGGGCGGAGAACGACGAGTCTCTGAGCACCGGGCGGTGCAAGG	7596
Qy	8242	CTACGACGCGTTTTCTGTCATCTCTGTGAGCAGTACCAGAACTGTCTCAGACAGCGCCAA	8301
Db	7597	GTACAGGAGATATTAGTCTCTCCGTGGAGCAATACCCAGAGCTTGTGACACAGTAGCAG	7656
Qy	8302	CAACATCCACTTCATGAGACACAGCGAGATGGGCCGAGGTTGACAGA	8348
Db	7657	CAACATCCAGTTTTTAAAGACAGAAATGAGATGGGAAAGAGGTAAACAA	7703

Search completed: June 25, 2004, 00:48:24
Job time : 2061 secs

This Page Blank (usp10)


```

; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3111 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-09-514-573-1

Query Match      20.5%; Score 1708.8; DB 4; Length 3111;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1713; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GTTTGTGGATGTGAGGAGCGCGGCGGAGCCATGAGCGTGAAGAGAGGAGAGCCCTTA 60
DB 300 GTTTGTGGATGTGAGGAGCGCGGCGGAGCCATGAGCGTGAAGAGAGGAGAGCCCTTA 359

QY 61 CCGCTCGTGAACCGCGGCGCGGAGCGCGGAGCGCGGTACACAGCTCGTCCCGGAGCAG 120
DB 360 CCGCTCGTGAACCGCGGCGCGGAGCGCGGAGCGCGGTACACAGCTCGTCCCGGAGCAG 419

QY 121 CGAGGAGGGAAGCCCGCGAGAAATCGTACAGCTCCAGCGAGACCCCTGAAGCGCTACGA 180
DB 420 CGAGGAGGGAAGCCCGCGAGAAATCGTACAGCTCCAGCGAGACCCCTGAAGCGCTACGA 479

QY 181 CCAGGAGCGCGCTAGCTATGCGAGCGCGGTCAAGGACATTTGTGCGGAGAGCGCGA 240
DB 480 CCAGGAGCGCGCTAGCTATGCGAGCGCGGTCAAGGACATTTGTGCGGAGAGCGCGA 539

QY 241 GGAATTTGCGGACAGGTGCGCACTTCAACCTCGGAGCGTGGGGCTGGAAGAGTAAC 300
DB 540 GGAATTTGCGGACAGGTGCGCACTTCAACCTCGGAGCGTGGGGCTGGAAGAGTAAC 599

QY 301 GCCCCTCACGGAGCCCTGTACCGGACAGACATTTGGCGTCCCAATGCGGCTACTCCAT 360
DB 600 GCCCCTCACGGAGCCCTGTACCGGACAGACATTTGGCGTCCCAATGCGGCTACTCCAT 659

QY 361 GGGGCTGCTGATGCGGACATGAGGCGTGAACCGTGTGCTCCCTCAGACACCCCGT 420
DB 660 GGGGCTGCTGATGCGGACATGAGGCGTGAACCGTGTGCTCCCTCAGACACCCCGT 719

QY 421 CGCTCTGCGGCGGAGCACACCGTCAAGGCGCAGCTCCTGCTGCTCCAGCGGCGCAA 480
DB 720 CGCTCTGCGGCGGAGCACACCGTCAAGGCGCAGCTCCTGCTGCTCCAGCGGCGCAA 779

QY 481 TTCCAAATCTACATCAACCGACACCGAGCATGAAGAACTGAGACTGATCATCGGCGCG 540
DB 780 TTCCAAATCTACATCAACCGACACCGAGCATGAAGAACTGAGACTGATCATCGGCGCG 839

QY 541 CCTCAGAACACCGCGGCTCCGAGCGCGCGCGCGCTCTCGACGCGCCACACCCC 600
DB 840 CCTCAGAACACCGCGGCTCCGAGCGCGCGCGCGCTCTCGACGCGCCACACCCC 899

QY 601 CAACAGACACACCGGCGCTCCATTAACCTCCCTGAACCGGCGCACTTCAAGCGGAGAG 660
DB 900 CAACAGACACACCGGCGCTCCATTAACCTCCCTGAACCGGCGCACTTCAAGCGGAGAG 959

QY 661 CAACCCGAGCGCGGCGCGGAGCGGAGCGGAGCGGAGCGGCGCGCGCGCG 720
DB 960 CAACCCGAGCGCGGCGCGGAGCGGAGCGGAGCGGAGCGGCGCGCGCGCG 1019

QY 721 CCAGGAGCTGCCACGCGCGGAGGAGTGGCTGCTCAAGCAACATCCCTCGGAGAG 780
DB 1020 CCAGGAGCTGCCACGCGCGGAGGAGTGGCTGCTCAAGCAACATCCCTCGGAGAG 1079

QY 781 CAGGAACCTAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840

```

RESULT 3

```

US-08-891-845-3
; Sequence 3, Application US/08991845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Slikowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:

```

```

DB 1080 CAGAAACCTTAGGCAAGCAGCCATTCTTAGGAGCATTTAGGAGCAACCTATTAGATGGA 1139
QY 841 CATTTCTCGGCGCTCCGCGCATGATGGGGCTTTACAGTGACGGGCACTTCTCTCTCAAGCC 900
DB 1140 CATTTCTCGGCGCTCCGCGCATGATGGGGCTTTACAGTGACGGGCACTTCTCTCTCAAGCC 1199
QY 901 TGGAGGCACTCTCCGCGCTCTTCTGCAACACATCATCACAGGTAACCATTCAGCTCCAGCAC 960
DB 1200 TGGAGGCACTCTCCGCGCTCTTCTGCAACACATCATCACAGGTAACCATTCAGCTCCAGCAC 1259
QY 961 AGTGTAATCTCTCGGCGCGGAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1020
DB 1260 AGTGTAATCTCTCGGCGCGGAGCGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1319
QY 1021 CTTCAAGAGCGCTCCAGTACTGTAAGTGAAGTGGCAGCGGAGCGGAGCGGAGCGGAGCGGAG 1080
DB 1320 CTTCAAGAGCGCTCCAGTACTGTAAGTGAAGTGGCAGCGGAGCGGAGCGGAGCGGAGCGGAG 1379
QY 1081 CTGAGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1380 CTGAGCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1439
QY 1141 CTGGGCACTGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1200
DB 1440 CTGGGCACTGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1499
QY 1201 TTGGGCTGTGCCAACCGGAGCTCTCCCTATACCCCTCAGGGGCACTGGCTTAGAGACCCC 1260
DB 1500 TTGGGCTGTGCCAACCGGAGCTCTCCCTATACCCCTCAGGGGCACTGGCTTAGAGACCCC 1559
QY 1261 TGACAGGAAAGCAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAA 1320
DB 1560 TGACAGGAAAGCAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAA 1619
QY 1321 TTTTCATAGATTCTGGAGAAATTTGATGGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
DB 1620 TTTTCATAGATTCTGGAGAAATTTGATGGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1679
QY 1381 CACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGTCATCTGAAATTCATGATGTGTC 1440
DB 1680 CACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTCTGTCATCTGAAATTCATGATGTGTC 1739
QY 1441 TCTGGGAAAGGAGCGGCTGGTGGGATTTATGGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
DB 1740 TCTGGGAAAGGAGCGGCTGGTGGGATTTATGGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1799
QY 1501 ACAGTTTGAATTTGTGGAGCTCTGGATGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
DB 1800 ACAGTTTGAATTTGTGGAGCTCTGGATGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1859
QY 1561 CCTAGAGGAGGAGCGGCGGAGCTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
DB 1860 CCTAGAGGAGGAGCGGCGGAGCTCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1919
QY 1621 CTTTCATCCAGTATTTGGATTCAGGAATCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
DB 1920 CTTTCATCCAGTATTTGGATTCAGGAATCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1979
QY 1681 GTTCAGAGTGGTTCCTTTCTCAGCACTGCCATTTAGTGC 1720
DB 1980 GTTCAGAGTGGTTCCTTTCTCAGCACTGCCATTTAGTGC 2019

```

; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WinPatIn (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/891,845
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/021640
 ; FILING DATE: 07/12/96
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M.
 ; REGISTRATION NUMBER: 40,378
 ; REFERENCE/DOCKET NUMBER: P1043
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1680 base pairs
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Single
 ; TOPOLOGY: Linear
 ; US-08-891-845-3

Query Match 20.0%; Score 1673.6; DB 3; Length 1680;

Best Local Similarity 99.8%; Pred. No. 0;
Matches 1676; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	575	CCGCCGCTCTGACGCCCCACACCCCAACACGACGACGCGGCTCCATTAACTCCCTG	634
DB	541	CCGCCGCTCTGACGCCCCACACCCCAACACGACGACGCGGCTCCATTAACTCCCTG	600
QY	635	AACCGGGCAACTTCACGCCGAGAGCAACCCACGAGCGGCGGCCACGAGCACTCGCTC	694
DB	601	AACCGGGCAACTTCACGCCGAGAGCAACCCACGAGCGGCGGCCACGAGCACTCGCTC	660
QY	695	TCCGAGAGCCCTCTGCGCGGCGGCCACGAGCTGCCACGCCAGGAGAAATGCGTG	754
DB	661	TCCGAGAGCCCTCTGCGCGGCGGCCACGAGCTGCCACGCCAGGAGAAATGCGTG	720
QY	755	CTCACAGCAATCCCTCGAGACGAGCAACCTAGCAGCAGGAGCAATTCCTAGGACA	814
DB	721	CTCACAGCAATCCCTCGAGACGAGCAACCTAGCAGCAGGAGCAATTCCTAGGACA	780
QY	815	TTGACAGCAACCTCAATTGAGATGACATTCCTCGCGGCTCCCGCCATGATGGGCTTAC	874
DB	781	TTGACAGCAACCTCAATTGAGATGACATTCCTCGCGGCTCCCGCCATGATGGGCTTAC	840
QY	875	AGTAGCGGCACTTCCTCTTCAAGCTGGAGGACCTCCCGCTCTTCTGACACCATCA	934
DB	841	AGTAGCGGCACTTCCTCTTCAAGCTGGAGGACCTCCCGCTCTTCTGACACCATCA	900
QY	935	CCAGGTACCCACTGACGTCAGACAGTGTACTCTCTCCGCGCCGACCCCTGCCCGC	994
DB	901	CCAGGTACCCACTGACGTCAGACAGTGTACTCTCTCCGCGCCGACCCCTGCCCGC	960
QY	995	AGCACCTTCGCGCGGCTTTAACTCAAGAGCCCTCAAGTACTGTAACTGGAAG	1054
DB	961	AGCACCTTCGCGCGGCTTTAACTCAAGAGCCCTCAAGTACTGTAACTGGAAG	1020
QY	1055	TGCGAGCCCTGAGCGCATCTCATCTCAGCCACTCTGGTCATCTCTGCGCATCTTT	1114
DB	1021	TGCGAGCCCTGAGCGCATCTCATCTCAGCCACTCTGGTCATCTCTGCGCATCTTT	1080
QY	1115	GTGGCATGCACTGTGTGGCTTAACTGGCACTGACGCGATGGAGGGGAGATGAT	1174
DB	1081	GTGGCATGCACTGTGTGGCTTAACTGGCACTGACGCGATGGAGGGGAGATGAT	1140
QY	1175	GAGTACGAGGACACAGCAGTGTGCTGCTGCAACCGGCTCTCCCTATACCC	1234
DB	1141	GAGTACGAGGACACAGCAGTGTGCTGCTGCAACCGGCTCTCCCTATACCC	1200
QY	1235	TCAGGGGCACTGGCTTAGACACCCCTGACAGGAAGGCAAGCAACAGAGGAAG	1294
DB	1201	TCAGGGGCACTGGCTTAGACACCCCTGACAGGAAGGCAAGCAACAGAGGAAG	1260
QY	1295	CCAGTAGTCTTCTTCAGAGGACAGTTTCAATGATTCAGAGAAATGATGGGAGG	1354
DB	1261	CCAGTAGTCTTCTTCAGAGGACAGTTTCAATGATTCAGAGAAATGATGGGAGG	1320
QY	1355	CGAGCTCCCAAGAGATTCCTCTGCGCACTTCTGGAGATCTCAAGTGTTCATAGCAT	1414
DB	1321	CGAGCTCCCAAGAGATTCCTCTGCGCACTTCTGGAGATCTCAAGTGTTCATAGCAT	1380
QY	1415	CCTGTGATCTGAAATTAATGTTCTCTGCGAAAGCAGCCCTGGTGGCATTTATGGC	1474
DB	1381	CCTGTGATCTGAAATTAATGTTCTCTGCGAAAGCAGCCCTGGTGGCATTTATGGC	1440
QY	1475	AGAAAGCCCTCCCTCTTACATACAGTTTCACTTGTGGAGCTCTGGATGGGAGG	1534
DB	1441	AGAAAGCCCTCCCTCTTACATACAGTTTCACTTGTGGAGCTCTGGATGGGAGG	1500
QY	1535	AGGTCTCTTAAACCCAGGAGCGGAGCTAGAGGGAGCCCGCCAGCTCTCGGGAACT	1594
DB	1501	AGGTCTCTTAAACCCAGGAGCGGAGCTAGAGGGAGCCCGCCAGCTCTCGGGAACT	1560
QY	1595	GTGCCCCCTCCAGCCATGACAGGCTTCATCCAGTATTTGGATTGAGAACTGACAC	1654
DB	1561	GTGCCCCCTCCAGCCATGACAGGCTTCATCCAGTATTTGGATTGAGAACTGACAC	1620

QY	35	ATGACGCTGAAGAGAGAGAGCTTACCGCTCGCTGACCGCGCGCGCGAGCGCGC	94
DB	1	ATGACGCTGAAGAGAGAGAGCTTACCGCTCGCTGACCGCGCGCGCGAGCGCGC	60
QY	95	CGCTACACCACTGCTCGCGGACAGCAGGAGGCAAGCCCGCAGAGAAATGTTACGC	154
DB	61	CGCTACACCACTGCTCGCGGACAGCAGGAGGCAAGCCCGCAGAGAAATGTTACGC	120
QY	155	TCCAGCAGACCCCTGAAGGCTACGACAGGAGCGCGCGCTAGCTATGGCAGCGCGTC	214
DB	121	TCCAGCAGACCCCTGAAGGCTACGACAGGAGCGCGCGCTAGCTATGGCAGCGCGTC	180
QY	215	AAGGACATTTGCGCGAGGAGCGAGAAATTCGCGGACAGGTCGCAACCTTCAACCTG	274
DB	181	AAGGACATTTGCGCGAGGAGCGAGAAATTCGCGGACAGGTCGCAACCTTCAACCTG	240
QY	275	CGGAGCTGGGGCTGGAGAGTAACGCCCTCAGGGACCCCTGACCGGACAGACAT	334
DB	241	CGGAGCTGGGGCTGGAGAGTAACGCCCTCAGGGACCCCTGACCGGACAGACAT	300
QY	335	GGCTGCCCAATGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC	394
DB	301	GGCTGCCCAATGCGGCTACTCCATGGGGGCTGGCTCTGATGCCGACATGGAGGCTGAC	360
QY	395	ACGGTGTGCTCCCTGAGCACCCCGTCTGTGGGGCGGAGACACAGGTCAGGGCGC	454
DB	361	ACGGTGTGCTCCCTGAGCACCCCGTCTGTGGGGCGGAGACACAGGTCAGGGCGC	420
QY	455	AGTCTCTGCTGTCCAGCGGGCAATTCACATCTCACATCTCACACCGAGCATGAA	514
DB	421	AGTCTCTGCTGTCCAGCGGGCAATTCACATCTCACATCTCACACCGAGCATGAA	480
QY	515	AACACTGAGACTGATCATCGGGGGGCTCGAGAAACACGCGGGCTTCGGAGCGCGCG	574
DB	481	AACACTGAGACTGATCATCGGGGGGCTTCAGAAACACGCGGGCTTCGGAGCGCGCG	540

QY 1655 TTGGCTTTTACAAATGACGGAAGAGTCAAGAGTGGTTTCCCTTCTCACCACAGCAATT 1714
 Db 1621 TTGGCTTTTACAAATGACGGAAGAGTCAAGAGTGGTTTCCCTTCTCACCACAGCAATT 1680

RESULT 4

US-09-514-573-3
 ; Sequence 3, Application US/09514573
 ; Patent No. 6500941
 ; GENERAL INFORMATION:
 ; APPLICANT: Schaefer, Gabriele M.
 ; APPLICANT: Sliwowski, Mark
 ; TITLE OF INVENTION: Gamma-Herregulin
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WinPatIn (Genentech)
 ; CURRENT APPLICATION DATA: US/09/514,573
 ; FILING DATE: 28 FEB 2000
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/891845
 ; FILING DATE: 10 JULY 1997
 ; APPLICATION NUMBER: 60/021640
 ; FILING DATE: 07/12/96
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M.
 ; REGISTRATION NUMBER: 40,378
 ; REFERENCE/DOCKET NUMBER: P1043
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1680 base pairs
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Single
 ; TOPOLOGY: Linear
 ; US-09-514-573-3

Query Match 20.0%; Score 1673.6; DB 4; Length 1680;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1676; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 35 ATGACGTGAAGGAGAGAGCCCTTACCGTGTGACCCCGCGCGCGCGAGCGC 94
 Db 1 ATGACGTGAAGGAGAGAGCCCTTACCGTGTGACCCCGCGCGCGCGAGCGC 60
 QY 95 CGCTACACCACTCGTCCGCGGACAGGAGGCGCAAGCCCGGAGAAATCGTACAGC 154
 Db 61 CGCTACACCACTCGTCCGCGGACAGGAGGCGCAAGCCCGGAGAAATCGTACAGC 120
 QY 155 TCCAGCGAGACCCCTGAAGGGCTTACGACAGGAGCCCGCCCTAGCCTATGGCAGCCGCGTC 214
 Db 121 TCCAGCGAGACCCCTGAAGGGCTTACGACAGGAGCCCGCCCTAGCCTATGGCAGCCGCGTC 180
 QY 215 AAGACATTTGCCGCGAGGAGGCGGAGGAAATTCGCGGACAGGTGCGCAACTTCACCGTC 274
 Db 181 AAGACATTTGCCGCGAGGAGGCGGAGGAAATTCGCGGACAGGTGCGCAACTTCACCGTC 240
 QY 275 CGGAGCTGGGGTGGGAAGTAACGCCCCCTCACGGGAGCCCTGTACGGGACAGCAATT 334

Db 241 CGGAGAGCTGGGGCTCGAAGAAAGTAAGGCCCTCTACGGGACCTGTACCGGACAGACATT 300
 QY 335 GGCCTGCCCCCAATATGCGGTACTTCCATGGGGGTGGCTCTGTATGCCGACATGGAGGCTGAC 394
 Db 301 GGCCTGCCCCCACTGCGGTACTTCCATGGGGGTGGCTCTGTATGCCGACATGGAGGCTGAC 360
 QY 395 ACGGTGCTGTCCCTGTGAGACACCCCGTGTGTGGGGCGGAGACACAGGTGAGGGCGC 454
 Db 361 ACGGTGCTGTCCCTGTGAGACACCCCGTGTGTGGGGCGGAGACACAGGTGAGGGCGC 420
 QY 455 AGCTCTCTGCTGTCCAGCGGGCCAAATTCCAATCTCACACTCACCGACACCGAGATGAA 514
 Db 421 AGCTCTCTGCTGTCCAGCGGGCCAAATTCCAATCTCACACTCACCGACACCGAGATGAA 480
 QY 515 AACACTGAGACTGATCATCCGGGGGCTGTGAGAAACACGGCGGCTCCGACGCGCGC 574
 Db 481 AACACTGAGACTGATCATCCGGGGGCTGTGAGAAACACGGCGGCTCCGACGCGCGC 540
 QY 575 CGCGCGCTCTCGACGCGCCACACCCCAACAGGACACAGCGGCTCCATTAATCTCCCTG 634
 Db 541 CGCGCGCTCTCGACGCGCCACACCCCAACAGGACACAGCGGCTCCATTAATCTCCCTG 600
 QY 635 AACCGGGGCAACTTCACGCGGAGGAGAACCCCGCGGCCCCCACGACCACTCGCTC 694
 Db 601 AACCGGGGCAACTTCACGCGGAGGAGAACCCCGCGGCCCCCACGACCACTCGCTC 660
 QY 695 TCCGAGAGGCGGCTGCGGGGGCGCCAGAGGCTGCGGCGGCTCCGACGAGGAGACTGGCTG 754
 Db 661 TCCGAGAGGCGGCTGCGGGGGCGCCAGAGGCTGCGGCGGCTCCGACGAGGAGACTGGCTG 720
 QY 755 CTCACAGCAACATCCCGCTGGAGACCAAGAACCTTAGGCAAGCAGCCATTCCTAGGGACA 814
 Db 721 CTCACAGCAACATCCCGCTGGAGACCAAGAACCTTAGGCAAGCAGCCATTCCTAGGGACA 780
 QY 815 TTGAGGACACCTCATTTAGATGGAATTTCTGCGGCTGCGGCGGCTCCGACCACTGAGGGCTTAC 874
 Db 781 TTGAGGACACCTCATTTAGATGGAATTTCTGCGGCTGCGGCGGCTCCGACCACTGAGGGCTTAC 840
 QY 875 AGTGAGGGGCACTTCTCTCAAGCTGAGGACCTCCCGCTCTTCTGACCAACATCA 934
 Db 841 AGTGAGGGGCACTTCTCTCAAGCTGAGGACCTCCCGCTCTTCTGACCAACATCA 900
 QY 935 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTTCTCCGCGCGCGGACCCCTGCCCCGC 994
 Db 901 CCAGGGTACCCACTGACGTCCAGCACAGTGTACTTCTCCGCGCGCGGACCCCTGCCCCGC 960
 QY 995 AGCACCTTTGCGCGCGCGGCTTTAACTCAAGAGCCCTCAAGTACTGTAAGTGGAG 1054
 Db 961 AGCACCTTTGCGCGCGCGGCTTTAACTCAAGAGCCCTCAAGTACTGTAAGTGGAG 1020
 QY 1055 TCGCAGAGCCCTGAGCGCCATCGTCACTCTCAGCCACTCTGGTCACTCTGCTGGCATACTTT 1114
 Db 1021 TCGCAGAGCCCTGAGCGCCATCGTCACTCTCAGCCACTCTGGTCACTCTGCTGGCATACTTT 1080
 QY 1115 GTGGCCATGACCTTTTGGGCTAACTGGGACCTGTGACCGGATGGAGGGGAGATGTAT 1174
 Db 1081 GTGGCCATGACCTTTTGGGCTAACTGGGACCTGTGACCGGATGGAGGGGAGATGTAT 1140
 QY 1175 GAGATACGGGAGACACAGCGAGTTGGCTGTGCCAAGCCAGCTCTCCCTATACCCC 1234
 Db 1141 GAGATACGGGAGACACAGCGAGTTGGCTGTGCCAAGCCAGCTCTCCCTATACCCC 1200
 QY 1235 TCAGGGGCGACTGGCTTTAGAGACCCCTGACAGGAAAGGCAAGGAAACACAGAGGAAAG 1294
 Db 1201 TCAGGGGCGACTGGCTTTAGAGACCCCTGACAGGAAAGGCAAGGAAACACAGAGGAAAG 1260
 QY 1295 CCCAGTACTTTCTTCCAGAGGACAGTTTCTAGATTTCTGAGAAATTTGATGTGGGAGG 1354
 Db 1261 CCCAGTACTTTCTTCCAGAGGACAGTTTCTAGATTTCTGAGAAATTTGATGTGGGAGG 1320
 QY 1355 CGAGCTCTCCAGAGATTCCTCTCGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCAT 1414
 Db 1321 CGAGCTCTCCAGAGATTCCTCTCGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCAT 1380

QY 1415 CCTGTGATCTGAATTAATCAATGTGTCTCTGGGAAAGCAGCCCTGGTTGGCATTATATGCG 1474
DB 1381 CCTGTGATCTGAATTAATCAATGTGTCTCTGGGAAAGCAGCCCTGGTTGGCATTATATGCG 1440
QY 1475 AGAAAAGGCTCCCTCTTACATACACAGTTTGAATTTGTGAGCTGTGATGACAGG 1534
DB 1441 AGAAAAGGCTCCCTCTTACATACACAGTTTGAATTTGTGAGCTGTGATGACAGG 1500
QY 1535 AGGCTCTTAACCCAGGAGGCGGAGGCTAGAGGGACCCCGCCAGTCTCGGGAACT 1594
DB 1501 AGGCTCTTAACCCAGGAGGCGGAGGCTAGAGGGACCCCGCCAGTCTCGGGAACT 1560
QY 1595 GTGCCCCCTCCAGCCATGAGACAGGCTTCCATCCAGTATTTGATTCAGGAATCTGGCAC 1654
DB 1561 GTGCCCCCTCCAGCCATGAGACAGGCTTCCATCCAGTATTTGATTCAGGAATCTGGCAC 1620
QY 1655 TTGGCTTTTACATGACGGAAGGCTCAGAGTGGTTTCTTCTCACCCTGCAAT 1714
DB 1621 TTGGCTTTTACATGACGGAAGGCTCAGAGTGGTTTCTTCTCACCCTGCAAT 1680

RESULT 5

US-08-891-845-11
; Sequence 11, Application US/08891845
; Patent No. 6096973
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2387 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; US-08-891-845-11

Query Match 14.28; Score 1185.6; DB 3; Length 2387;
Best Local Similarity 98.8%; Pred. No. 1.5e-290;
Matches 1194; Conservative 0; Mismatches 14; Indels 0; Gaps 0;
QY 530 CATCCGGCGGCGCTTGAGAACACACGCGGCTCCGACGCGCGCGCGCTCTCGCAC 589
DB 1 CATCCGGCGGCGCTTGAGAACACACGCGGCTCCGACGCGCGCGCGCTCTCGCAC 60

QY 590 GCCCACACCCCAACACAGCAGCCGCGCTTCCATTAACTCCCTGAAACCGGGCACTTC 649
DB 61 GCCCACACCCCAACACAGCAGCCGCGCTTCCATTAACTCCCTGAAACCGGGCACTTC 120
QY 650 AGCGGAGGAGCAACCCAGCCCGGCCCCCAACGAGCACTCGCTCTCCGAGAGCCCCCT 709
DB 121 AGCGGAGGAGCAACCCAGCCCGGCCCCCAACGAGCACTCGCTCTCCGAGAGCCCCCT 180
QY 710 GCCGCGCGCGCCAGGAGCCTTGGCCACGCGCCAGGAGAACTGGCTGTCTCAACAGCAATC 769
DB 181 GCCGCGCGCGCCAGGAGCCTTGGCCACGCGCCAGGAGAACTGGCTGTCTCAACAGCAATC 240
QY 770 CCCTGGAGACCAAGAACTTAGGCAAGAGCCATTCTTAGGACATTGAGGACAACTC 829
DB 241 CCCTGGAGACCAAGAACTTAGGCAAGAGCCATTCTTAGGACATTGAGGACAACTC 300
QY 830 ATTGAGATGGACATCTCGGCGCTCCCGCCATGATGGGCTTACAGTACGGGCACTTC 889
DB 301 ATTGAGATGGACATCTCGGCGCTCCCGCCATGATGGGCTTACAGTACGGGCACTTC 360
QY 890 CTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCAATCACCAGGCTACCCACTG 949
DB 361 CTCTTCAAGCCTGGAGGACCTCCCGCTCTTCTGCACCAATCACCAGGCTACCCACTG 420
QY 950 AGTCCAGCACAGTGTACTCTCTCCGCGCGGACCTGCCCCGACAGCACTTCGCCCGG 1009
DB 421 AGTCCAGCACAGTGTACTCTCTCCGCGCGGACCTGCCCCGACAGCACTTCGCCCGG 480
QY 1010 CCGGCTTTTAACCTCAAGAGCCCTCCAAAGTACTGTAACTGGAAGTGGCAGCCCTGAGC 1069
DB 481 CCGGCTTTTAACCTCAAGAGCCCTCCAAAGTACTGTAACTGGAAGTGGCAGCCCTGAGC 540
QY 1070 GCCATGCTATCTCAGCCACTCTGGTCACTCTGTGTCGATCTTTGTCGATCTGACCTG 1129
DB 541 GCCATGCTATCTCAGCCACTCTGGTCACTCTGTGTCGATCTTTGTCGATCTGACCTG 600
QY 1130 TTTGCGCTTAACTGSCACCTGSCACCGATGGAGGGCAGATGTATGATGATCAGGAGGAC 1189
DB 601 TTTGCGCTTAACTGSCACCTGSCACCGATGGAGGGCAGATGTATGATGATCAGGAGGAC 660
QY 1190 ACAGCCAGCAGTTGGCTGTGSCAACCCAGCTCTCCCTATACCCCTCAGGGGCACTGGC 1249
DB 661 ACAGCCAGCAGTTGGCTGTGSCAACCCAGCTCTCCCTATACCCCTCAGGGGCACTGGC 720
QY 1250 TTAGAGACCCCTGACAGGAAGGCAAGAGAACACAGAGGAAGCCAGTATTTCTTT 1309
DB 721 TTAGAGACCCCTGACAGGAAGGCAAGAGAACACAGAGGAAGCCAGTATTTCTTT 780
QY 1310 CCAGAGGACAGTTTTCATAGATTCTGGAGAAATTTGATGTGGGAAGGCGAGCTCCAGAG 1369
DB 781 CCAGAGGACAGTTTTCATAGATTCTGGAGAAATTTGATGTGGGAAGGCGAGCTCCAGAG 840
QY 1370 ATTCTCTCTGCACTTTCTGAGATCTCAAGTGTTCATAGACCATCTCTGTCATCTGAAA 1429
DB 841 ATTCTCTCTGCACTTTCTGAGATCTCAAGTGTTCATAGACCATCTCTGTCATCTGAAA 900
QY 1430 TTCAATGTCTCTGGGAAGGCGAGCCCTGGTGGCATTTATGGCAAAAGGCGCTCCCT 1489
DB 901 TTCAATGTCTCTGGGAAGGCGAGCCCTGGTGGCATTTATGGCAAAAGGCGCTCCCT 960
QY 1490 CTTTCACATACACAGTTTGTGAGCTGTGATGGAGAGGCTCTTAACCCAG 1549
DB 961 CTTTCACATACACAGTTTGTGAGCTGTGATGGAGAGGCTCTTAACCCAG 1020
QY 1550 GAGCGCGGAGCCTAGAGGGGACCCCGCCAGTCTCGGGAACTGTGCCCCCTCCAGC 1609
DB 1021 GAGCGCGGAGCCTAGAGGGGACCCCGCCAGTCTCGGGAACTGTGCCCCCTCCAGC 1080
QY 1610 CATGAGACAGGCTTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTTGGTTTTACAAT 1669
DB 1081 CATGAGACAGGCTTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTTGGTTTTACAAT 1140


```

; Sequence 1072, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1072
; LENGTH: 447
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(447)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1072

Query Match 4.7%; Score 389; DB 4; Length 447;
Best Local Similarity 94.0%; Pred. No. 8e-89;
Matches 426; Conservative 0; Mismatches 0; Indels 27; Gaps 1;

QY 2341 CTGCGCGGCGGCAAGTGGAGTGCAGCCCTGGCTGGATGGGAAACACTGCACCATCGC 2400
Db 22 CTGCGCGGCGGCAAGTGGAGTGCAGCCCTGGCTGGATGGGAAACACTGCACCATC-- 79

QY 2401 TCACTATCTGGATAGGTAGTTAAAGAGGTTCCCTGGGTTGTCAATGGCAACGGCAG 2460
Db 80 -----GAGGGTTGCTGCTGGGTTGTGCAATGGCAACGGCAG 114

QY 2461 ATGTACCTTAGACCTGAATGGTGGCACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 2520
Db 115 ATGTACCTTAGACCTGAATGGTGGCACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 174

QY 2521 CTGTGACACTTCCATGGAGACTGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2580
Db 175 CTGTGACACTTCCATGGAGACTGCCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 234

QY 2581 GGTGGAGTGCATGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Db 235 GGTGGAGTGCATGACGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 294

QY 2641 CCTGGCTCCCTTAACCTCTGGACATCATCCAGGAGACACAGTCCCTGTGTGCACAGCA 2700
Db 295 CCTGGCTCCCTTAACCTCTGGACATCATCCAGGAGACACAGTCCCTGTGTGCACAGCA 354

QY 2701 GAACCTACACTCTTCTTATGACCGCATCAAGTTCCTGCTGGGAGGAGACAGCAGCAGAT 2760
Db 355 GAACCTACACTCTTCTTATGACCGCATCAAGTTCCTGCTGGGAGGAGACAGCAGCAGAT 414

QY 2761 AATCCCGGGGAGAACCCCTTTGATGAGGGCA 2793
Db 415 AATCCCGGGGAGAACCCCTTTGATGAGGGCA 447

RESULT 8
US-09-594-407/c
; Sequence 407, Application US/0976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143

```

```

; SOFTWARE: PERL Program
; SEQ ID NO 407
; LENGTH: 2007
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 474200.1
US-09-976-594-407

Query Match 2.5%; Score 208.6; DB 4; Length 2007;
Best Local Similarity 72.7%; Pred. No. 1.1e-42;
Matches 296; Conservative 0; Mismatches 109; Indels 2; Gaps 2;

QY 7941 GGGGGCGGCAACCTCTGAGATGGGTCAACGTCACCTGTGTCCAGATCAACACAGTAC 8000
Db 2007 CGGGCCGAAGCGCTGGAGAACGGCATCAACGTGACGGTGTGCGAGTCCACACGGTGG 1948

QY 8001 TTAATGGCAGGACTAGACGCTTACACAGACATCCAGTCCAGTACGGGGACT-GTGCTTG 8059
Db 1947 TGATCGGCGAGCAGCGCAGGTTCCGGACGCTGGAGATGCAGTTCGGCGCGCTGCGCGTGC 1888

QY 8060 AACACAGCTACGGGACAACTTGGATG-AGGAGAGGCGAGGTCCTGGAGCTGGCGCG 8118
Db 1887 CACGTGGCTACGGCATGACCTTGGAGAGGAGAGGCGCGCATCTCTGAGCAGGGCGCG 1828

QY 8119 GCAGAGAGCCGTGGCGCAAGCGTGGCGCCGCGAGCAGCAGAGACTGCGGGAAGGGAGGA 8178
Db 1827 CGAGCGCGCGCTCGCCCGGGCTGGCGCGCGAGCAGCAGCGCGTGGCGAGGAGGA 1768

QY 8179 AGGCTGGCGGCTTGGACAGAGGGGAGAGCAGCAGGTCGTGAGCAGAGGGGGGTGCA 8238
Db 1767 GGGCGCGCGCTTGGACAGAGGGGAGAGCAGCAGCAGCTGCTGAGCGCGCGCAAGGTGCA 1708

QY 8239 AGGCTACAGCGGCTTTTTCGTGATCTCTGTGAGCAGTACCCAGAACTGTGACAGCGC 8298
Db 1707 GGGCTACAGCGGCTTACGTACTCTGCTGGAGCAGTACCCGAGCTGCGCGAGCGC 1648

QY 8299 CAACAACATCCACTTCAATGAGACAGCGAGATGGCGGAGGTGAC 8345
Db 1647 CAACAACATCCAGTTCCTGCGGAGAGCGAGATCGCGAGGAGTAA 1601

RESULT 9
US-09-620-312D-75
; Sequence 75, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/489,725

```


RESULT 11
PCT-US95-11684-3
; Sequence 3, Application PC/TUS9511684
; GENERAL INFORMATION:
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE
; ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
; TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10665 North Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11684
; FILING DATE: 14-SEP-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,359
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Logan, April C.
; REGISTRATION NUMBER: 33,950
; REFERENCE/DOCKET NUMBER: BEC0019P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6049 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 309..5740
PCT-US95-11684-3

Query Match 1.1%; Score 92.2; DB 5; Length 6049;
Best Local Similarity 49.9%; Pred. No. 7.4e-13;
Matches 331; Conservative 0; Mismatches 308; Indels 24; Gaps 3;
QY 1730 TGCCCCAGCAACTCTATGCAATGGTGAAGTCTCTGGACCTGCCACTGCTTCTCTG 1789
DB 1254 TGCCCCAATGACTGCTTTGACCGGGCGCTGATCAATGGGACCTTCTTCTCGAGGAG 1313
QY 1790 GGTTCCTGGCCCCGACTGTGGCAGAGCCTCTGCCCCGCTGTGTAGCGGAATGCG 1849
DB 1314 GGCTACACTGGAGAGGACTGGCGAGCTGACCTGCCGCCAACAACTGCAACGGCAACGGG 1373
QY 1850 CAATACATGAAGCGACAGTCTTGCCACAGTGGCTGGAAAGCGCTGAGTGCATGTG 1909
DB 1374 CGCTGTGAGAACGGGCTGTGTGTGCTATGAGGCTTCTGGGGGATGACTG----- 1426
QY 1910 CCCACCAACCAAGTGTATCGATGTGGCTGCAAGCAACATGCAACCTGCAATCAGGGCACC 1969
DB 1427 --CAGCCAGAAGAGTGTCCCGAAGACGTGCAATAACCGCGGGCGCTGCGTGGATGGCGC 1484
QY 1970 TGATCTGCAACCTGCTCAAGGGCGAGAGTGTGAGGAAGTGAAGTCAATGAGACCCC 2029
DB 1485 TGTGTGTGCAATGAGGGGTACCTGGGGGAGGACTGTGGGGAGCTGCGGTGCCCAACGAC 1544
QY 2030 ACATGTTTCAGCGCCGGGGTGTCTGCGTGAGAGGGCAATGCCATGCTTTTGTGGGATGGGA 2089

DB 1545 ---TGCCCAACCCGGCGCTGTCATCAACCGGCGAGTGTGTGTGATGAGGATTCATT 1601
QY 2090 GGCACCAACTGCGAGACCCCCAGGGGCCACATGCTTAGACCAAGTGTTCAGGCCACGGAACC 2149
DB 1602 GGGGAGGACTGCGAGAGCTCGGTGCCCAACACTGCCAGCAACCGCGGCGCTGCAT- 1660
QY 2150 TTCCTCCCGACACCGGCTTTGAGAGTGTGACCCAAAGCTGGAGTGGACACAGCTGTCT 2209
DB 1661 -----CAATGGGCGAGTGGCGAGTGCACAGAGGATTCATCGGGGAGGACTGCGGG 1709
QY 2210 ATCGAGATCTGTGTCGCACTGTGCTGGCCATGGCGTGTGCGTAGGGGGCACCCTGCCGC 2269
DB 1710 GAGCTGGGTGTCCCAAGACTGCAACAGCATGGGCGCTGCGTCAATGGCAGTGGTG 1769
QY 2270 TCGAGAGTGTGATGGGGCGAGCTGTGCAACAGCGGGCTGCCACCCCGCTGTGCC 2329
DB 1770 TGTGATGAGGGGTACACAGGGGAGGACTGCGGGGAGTTGGCGTCCCCCAACGACTGCCAC 1829
QY 2330 GAGCATGGGACCTGCCGCGAGCGCAAGTGGAGTGCAGCCCTGGCTGGAATGGCGAACAC 2389
DB 1830 AACCGCGGCGCTGCTGGAGGAGCGCTGTGTGTGACAAAGGCTTCATCGGGGAGGAC 1889
QY 2390 TGC 2392
DB 1890 TGC 1892
RESULT 12
US-08-793-273C-1
; Sequence 1, Application US/08793273C
; Patent No. 6482410
; GENERAL INFORMATION:
; APPLICANT: Crossin, Kathryn L.
; APPLICANT: Phillips, Greg
; APPLICANT: Prieto, Anne L.
; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE ATTACHMENT AND
; FILE REFERENCE: BEC0022S
; CURRENT APPLICATION NUMBER: US/08/793,273C
; CURRENT FILING DATE: 1997-05-22
; PRIOR APPLICATION NUMBER: PCT/US95/11684
; PRIOR FILING DATE: 1995-09-14
; PRIOR APPLICATION NUMBER: 08/308,359
; PRIOR FILING DATE: 1994-09-16
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7286
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (55)...(6654)
US-08-793-273C-1

Query Match 1.0%; Score 84; DB 4; Length 7286;
Best Local Similarity 49.1%; Pred. No. 1e-10;
Matches 327; Conservative 0; Mismatches 315; Indels 24; Gaps 3;
QY 1730 TGCCCCAGCAACTCTATGCAATGGTGAAGTCTCTGGACCTGCCACTGCTTCTCTG 1789
DB 715 TGCCCCAGGACTGCAATGATGACCGGCAAGTGGCTGTGATGAGTCTGTCATCTGTTTGAA 774
QY 1790 GGTTCCTGGCCCCGACTGTGGCAGAGCCTCTGCCCCGCTGTCTGTAGCGGAATGCG 1849
DB 775 GGCTACGCGGCTGACTGCGAGCGGTGAAATCTGCCAGTGCCTGCACTGAGGAGCACGGC 834
QY 1850 CAATACATGAAGCGACAGTCTTGCCACAGTGGCTGGAAGCGCTGAGTGCATGTG 1909
DB 835 ACATGTGTAGATGCTTGTGTGTGTCACAGATGGCTTTTCAGGCGGATGACTGCAACAAG 894
QY 1910 CCCACCAACCAAGTGTATCGATGTGGCTGCGCAACCAATGCCACTGCAATCAGGGCACC 1969

```
Db 895 CCTCTGTCTCAACAAT-----TGCTAACACCGTGGAGCATGCGTGGGAATGAG 945
Qy 1970 TGCATCTGCAACCCCTGCTACAAGGCGGAGAGCTGTGAGGAAGTGGACTCATGGACCCC 2029
Db 946 TGCCTGTGTATGAGGGTTTCACGGGGAAGACTGCACTGAGCTCATCTGCCCAATGAC 1005
Qy 2030 ACATGTTTCAGCCCGGGGTGTCTGCTGAGAGGGAATGCCAATGCTTTTGGGATGGGA 2089
Db 1006 ---TGCTTTCAGCCCGGGCGGTGATCAATGGCACTGCTACTGCGAAGAAGGCTTCACA 1062
Qy 2090 GGCACCAACTGCGAGAGCCCCAGGCGCACATGCTTAGACCAAGTGTTCAGGCCACGGAACC 2149
Db 1063 GGTGAAGACTGCGGGAACCACTGCCCA-----CATGCTGCCAACCAG 1110
Qy 2150 TTCTCCCGGACACCGGGCTTTCAGCTGTGACCAAGCTGGAGTGGACACGACTGTCT 2209
Db 1111 GGCCTGTGTAGGAGGGGCACTGTGTATGATGAGGGCTTTCGGGTGGGACTGCAGC 1170
Qy 2210 ATCAGATCTGTGCTGCGACTGTGTGGCCATGGCTGTGCTAGGGGSCACTGCCGC 2269
Db 1171 GAGAAGAGGTGCTCTGCTGACTGTCACAAATCGTGGCGCTGTGTAGACGGCGGTGTAG 1230
Qy 2270 TGCAGGATGGCTGGAGCGGCGAGCTGCGACACGAGCGGCGCTGCCACCGCGCTGTGCC 2329
Db 1231 TGTGATGATGTTTCACTGGAGCTGACTGTGGGAGCTCAAGTGTCCCAATGCTGCAGT 1290
Qy 2330 GAGATGGGACCTCCCGGACCGCAAGTGGAGTGGAGCGCCCTGGCTGGAATGGGAACAC 2389
Db 1291 GGCATGGCGCTGTGTAATGGGCACTGTGTGTGATGAGGGCTATACCTGGGAGGAC 1350
Qy 2390 TGCACC 2395
Db 1351 TGCAGC 1356
```

RESULT 13

```
PCT-US95-11684-1
; Sequence 1, Application PC/TUS9511684
; GENERAL INFORMATION:
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE
; TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 North Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/11684
; FILING DATE: 14-SEP-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/308,359
; FILING DATE: 16-SEP-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Logan, April C.
; REGISTRATION NUMBER: 33,950
; REFERENCE/DOCKET NUMBER: BEC0019P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 1:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7286 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 55..6654
; OTHER INFORMATION: /product= "cytotactin"
PCT-US95-11684-1
```

```
Query Match 1.0%; Score 84; DB 5; Length 7286;
Best Local Similarity 49.1%; Pred. No. 1e-10;
Matches 327; Conservative 0; Mismatches 315; Indels 24; Gaps 3;
Qy 1730 TGCCCCCAGCAACTGCTATGGCAATGTCAGCTGCAFTCTCTGGGACCTGCCACTGCTTCCTG 1789
Db 715 TGCCCCCAGCGACTGCAATGACCCAGGCAAGTGCCTGGAATGGAGTCTGCATCTGTTTCGAA 774
Qy 1790 GGTTCCTGGGCCCCGACTGTGGCAGAGCTTCTGCCCGTGCTCTGTAGCGGAATGCC 1849
Db 775 GGCTACGGCGGTGACTGCGAGCGGTGAATCTGCCCAGTGCCCTGCGAGTGAGGACCGC 834
Qy 1850 CAATACATCAAAAGGAGAGATGCTTGTGCCACAGTGGCTGAAAAGGCGCTGAGTGCATGTG 1909
Db 835 ACATGTGTAGTGGCTTGTGTGTGTGCCACAGTGGCTTTCGAGGCGATGACTGCAACAAG 894
Qy 1910 CCCACCAACCACTGTATGATGTGCTGCGCTGCGAGCAACCATGGCACCTGCATCAACGGGACC 1969
Db 895 CTTCTGTCTCAACAAT-----TGCTACAAACCGTGGACGATGCTGGAGAAATGAG 945
Qy 1970 TGCATCTGCAACCCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCCC 2029
Db 946 TCGCTGTGTGATGAGGGTTTACGGGCGAAGTCTGAGTGAAGTCTCATCTGCCCAATGAC 1005
Qy 2030 ACATGTTTCAGCCCGGGGTGTCTGCTGAGAGCGCAATGCCAFTGCTTTTGGGATGGGA 2089
Db 1006 ---TGCTTCGACCGGGCGCTGTCAATCAATGGCACCTGTCTACTGCGAAGAAGGCTTCACA 1062
Qy 2090 GGCACCAACTGCGAGACCCCGAGGCGCACATGCTTAGACCAGTGTTCAGGCCACGGAACC 2149
Db 1063 GGTGAAGACTGCGGGAACCCACCTTGCCCA-----CATGCTTGCCACACCCAG 1110
Qy 2150 TTCTCTCCCGGACACCGGGCTTTGCACTGTGACCCCAAGCTGGACTGGACACGACTGTTCT 2209
Db 1111 GCGCGGTGTGAGGAGGGGCACTGTGTATGTGATGAGGGCTTTGCGGCTGTGACTGCAGC 1170
Qy 2210 ATCGAGATCTGTCTGCCGACTGTGGTGGCCATGGCGTGTGCTAGGGGGGCACTGCCGC 2269
Db 1171 GAGAAGAGGTGCTCTGCTGACTGTCAAAATCGTGGCGCTGTGTAGACGGGCGGTGTGAG 1230
Qy 2270 TCGGAGGATGGCTGGATGGGCGGAGCTTCCGACACGAGCGGCGCTGCCACCGGCTGTGCC 2329
Db 1231 TGTGATGATGTTTCACTGGAGTGAATGCTGGGAGCTCAAGTGTCCCAATGGCTGCAGT 1290
Qy 2330 GAGCATGGGACCTGCGCGACCGCAAGTGGCAGTGCAGCCCTGGCTGGGAATGGCAACAC 2389
Db 1291 GGCCATGGCGCTGTGTCATGGGCACTGTGTGTGTGATGAGGGCTATACCTGGGAGGAC 1350
Qy 2390 TGCACC 2395
Db 1351 TGCAGC 1356
```

RESULT 14

```
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
```

;; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
;; NUMBER OF SEQUENCES: 52
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Foley & Lardner
;; STREET: 1800 Diagonal Road, Suite 500
;; CITY: Alexandria
;; STATE: VA
;; COUNTRY: USA
;; ZIP: 22313-0299
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; OPERATING SYSTEM: IBM PC compatible
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/232.463
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US/07/935.313
;; FILING DATE:
;; APPLICATION NUMBER: EP 91 114 300.6
;; FILING DATE: 26-AUG-1991
;; ATTORNEY/AGENT INFORMATION:
;; NAME: BENT, Stephen A.
;; REGISTRATION NUMBER: 29,768
;; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703)836-9300
;; TELEFAX: (703)683-4109
;; TELEX: 899149
;; INFORMATION FOR SEQ ID NO: 14:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 7218 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match 0.7%; Score 58.4; DB 1; Length 7218;

Best Local Similarity 4.4%; Pred. No. 0.00031;

Matches 17; Conservative 220; Mismatches 151; Indels 0; Gaps 0;

Qy	7894	GAAACAGGACCTTCAGAGGTGACCTGCGCATCTGGGCTCAGTGGGGGGCGGGAAC	7953
Db	1427	RR	1368
Qy	7954	CTTGAGAGATGGGTCAACGTCACTGTGTCCAGATCAACACAGTACTTAATGGCAGAC	8013
Db	1367	RR	1308
Qy	8014	TAGACGTACACAGATCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG	8073
Db	1307	RR	1248
Qy	8074	GACACCTTGATGAGAGAGGACGAGTCTGAGCTGGCGCGGAGAGAGAGAGAGAGAGAG	8133
Db	1247	RR	1188
Qy	8134	CCAAAGCTGGGCGGAG	8193
Db	1187	RR	1128
Qy	8194	GACAGAGGAG	8253
Db	1127	RR	1069
Qy	8254	TTTCGTGATCTCTGTCGAGCAGTACCA	8281
Db	1067	RATCGCAAGCTCCCTCGACCTCGAGCA	1040

RESULT 15
US-08-404-665-3
; Sequence 3, Application US/08404665
; Patent No. 5591583
; GENERAL INFORMATION:
; APPLICANT: Reid, Robert A.
; APPLICANT: Ackley, Rhonda L.
; APPLICANT: Hemperly, John J.
; TITLE OF INVENTION: HUMAN RESTRICTIN AND NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
; ADDRESSEE: Company
; STREET: 1 Becton Drive
; CITY: Franklin Lakes
; STATE: NJ
; COUNTRY: US
; ZIP: 07417
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/404.665
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fugit, Donna R.
; REGISTRATION NUMBER: 32,135
; REFERENCE/DOCKET NUMBER: P-3341
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4724 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-404-665-3

Query Match 0.7%; Score 57; DB 1; Length 4724;

Best Local Similarity 50.9%; Pred. No. 0.00055;

Matches 135; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Qy	2129	CAGTGTTCAGCCACGGAACCTTCCTCCCGGACACCGGGCTTTGACAGCTGTGACCAAGC	2188
Db	985	CAGTGTTCAGCCACGGAACCTTCCTCCCGGACACCGGGCTTTGACAGCTGTGACCAAGC	1044
Qy	2189	TGGACTGACACGAGCTGTTCATCGAGATCTGTCTCCGACTGTGTGGCCATGGCGGTG	2248
Db	1045	TGTTTGGCAAGATTGCTCGAGCCCTACTGCCCGCTGGTGTCTCCAGCGGGGGGTG	1104
Qy	2249	TGGTAGGGGGGCACTGCGCGCTCGAGAGTGGCTGGATGGGGGGGAGCCTTGGACCAAGCGG	2308
Db	1105	TGTGTGATGGCCAGTCATCTGTGACAGGAGTACAGCGGGGATGACTGTTCGAACTC	1164
Qy	2309	GCCTGCCACCGCGCTGTGCCGAGCTGGGACCTGCCGACGCAAGTGGAGTGCAGC	2368
Db	1165	CGTGGCCCAACAGACTGAGCTCCCGGGGCTCTGCTGACGGGAGTGTGTCTGTGAA	1224
Qy	2369	CCTGGCTGGAATGCGCAACTGCA	2393
Db	1225	GAGCCCTACACTGGCGAGGACTGCA	1249

Search completed: June 25, 2004, 00:13:59

Job time : 383 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 17:14:23 ; Search time 18786 Seconds
(without alignments)
13279.498 Million cell updates/sec

Title: US-10-029-020-13

Perfect score: 8354

Sequence: 1 gtttggtggtggtgagagc.....ccggaggtcacagagagac 8354

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthm:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_hic:*

9: gb_estl:*

10: gb_est2:*

11: gb_hic:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rtd:*

26: em_gss_phg:*

27: em_gss_vrt:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5857.4	70.1	6246	29 AY413475	AY413475 Homo sapi
2	5045.8	60.4	5970	29 AY413476	AY413476 Pan trogl
3	2353	28.2	5094	29 AY405420	AY405420 Homo sapi
4	2347.8	28.1	3190	29 AY413477	AY413477 Mus muscu

5	2282.2	27.3	5069	29 AY405421	AY405421 Pan trogl
6	2232.8	26.7	5087	29 AY405422	AY405422 Mus muscu
7	1407.8	16.9	3038	11 AK037897	AK037897 Mus muscu
8	743.8	8.9	1284	11 AK039472	AK039472 Mus muscu
9	713.8	8.5	718	9 AL045768	AL045768 DKF2p234F
10	695.2	8.3	856	14 CF534506	CF534506 UI-M-G10-
11	630.8	7.6	827	12 BM947883	BM947883 UI-M-BG0P
12	629	7.5	870	10 BG036207	BG036207 602326960
13	613.6	7.3	728	14 CF536185	CF536185 UI-M-G10-
14	609.2	7.3	770	14 CB524534	CB524534 UI-M-P10-
15	602.6	7.2	2627	11 AK050784	AK050784 Mus muscu
16	581	7.0	1031	29 CNS02PHB	AL207992 Tetraodon
17	580.8	7.0	728	14 CB520016	CB520016 UI-M-G10-
18	567.2	6.8	572	13 BX504178	BX504178 DKF2p86D
19	561.4	6.7	668	13 BO563239	BO563239 gi01b09.Y
20	544.8	6.5	694	14 CF723353	CF723353 UI-M-GV0-
21	534.6	6.4	637	14 CF171639	CF171639 B0845C10-
22	531.2	6.4	798	10 BF144774	BF144774 601791454
23	528.6	6.3	632	10 BB655038	BB655038 BB655038
24	525	6.3	621	14 CF182790	CF182790 UI-M-EY0-
25	518.6	6.2	525	9 AL046228	AL046228 DKF2p434E
26	514	6.2	592	10 BF349559	BF349559 MRO-HT015
27	497	5.9	693	13 EQ210132	BQ210132 UI-R-EP0-
28	494.2	5.9	620	14 CB532612	CB532612 757081 MA
29	491	5.9	620	14 CB532245	CB532245 756697 MA
30	489.4	5.9	566	13 BU072782	BU072782 im47902.Y
31	486	5.8	900	13 BU267121	BU267121 603506284
32	485.8	5.8	894	13 BU151768	BU151768 AGENCOURT
33	473.2	5.7	545	10 BF075317	BF075317 224264 MA
34	473	5.7	734	13 BU233488	BU233488 603408464
35	471.4	5.6	675	10 BE374675	BE374675 601224176
36	469.2	5.6	603	10 BB611598	BB611598 BB611598
37	468.6	5.6	574	9 AV589581	AV589581 AV589581
38	464	5.6	600	14 CA527167	CA527167 8032-72 M
39	456.2	5.5	868	29 CNS03M3H	AL263222 Tetraodon
40	455.2	5.4	568	13 BU214946	BU214946 603106244
41	451.6	5.4	724	13 BU387957	BU387957 603583120
42	449.8	5.4	819	13 BQ443891	BQ443891 UI-M-EW0-
43	449.8	5.4	951	13 BU839812	BU839812 AGENCOURT
44	446.8	5.3	757	10 BE741784	BE741784 601595546
45	443.4	5.3	461	10 BF932388	BF932388 IL2-NT019

ALIGNMENTS

RESULT 1

AY413475

LOCUS

DEFINITION

AY413475

Genomic survey sequence.

AY413475.1 GI:39769437

GSS.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Homo sapiens

REFERENCE

1 (bases 1 to 6246)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,

Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Adams,M.D. and Cargill,M.

Inferred nonneutral evolution from human-chimp-mouse orthologous

gene trios

Science 302 (5652), 1960-1963 (2003)

JOURNAL

PUBMED

14671302

REFERENCE

2 (bases 1 to 6246)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,

Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Adams,M.D. and Cargill,M.

Direct Submission

AY413475 6246 bp DNA linear GSS 12-DEC-2003

Homo sapiens HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,

Genomic survey sequence.

AY413475.1 GI:39769437

GSS.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Homo sapiens

Homo sapiens

REFERENCE

1 (bases 1 to 6246)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,

Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Adams,M.D. and Cargill,M.

Inferred nonneutral evolution from human-chimp-mouse orthologous

gene trios

Science 302 (5652), 1960-1963 (2003)

JOURNAL

PUBMED

14671302

REFERENCE

2 (bases 1 to 6246)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,

Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,

Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,

Adams,M.D. and Cargill,M.

Direct Submission

JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	Location/Qualifiers				
source	1..6246				
gene	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	<1..>6246				
ORIGIN	/locus_tag="HCM4903"				
Query Match	70.1%; Score 5857.4; DB 29; Length 6246;				
Best Local Similarity	95.4%; Pred. No. 0;				
Matches 6006; Conservative	0; Mismatches 233; Indels 57; Gaps 4;				
QY	2055	TGAGAGCGAATGCCATTGCTTTGTGGATGGGAGGACCACTGCGAGACCCCGAGG	2114		
DB	2	TGAAGGCGAGTGTGTGCCACAGTGTGGAAAGCGCTGAGTGGATGTGCCACCA	61		
QY	2115	CCACATGCTTAGACCAAGTGTAGGCCACGGAACCTTCTCCCGGACACCGGGCTT	2174		
DB	62	ACCAAGTGTATCGATGTGGCTGTGAGCAACCATGGCACTGC--ATCAGGGCACCT	118		
QY	2175	GCTGTGACCCAGTGAAGTGGACACGACTGTTCTATCGAGATCTGTGCTGCCACT	2234		
DB	119	TCTGCAACCTTGGCTACAAGGCGAGAGCTGTGAGGAAGAGATCTGTGCTGCC	178		
QY	2235	GTGCCCATGGCTGTGCTGAGGGGACCTGCGCGTGCAGGAGTGGTGGATGGGG	2294		
DB	179	GTGCCCATGGCTGTGCTGAGGGGACCTGCGCGTGCAGGAGTGGTGGATGGGG	238		
QY	2295	CCTCGACACCGGGCTGCCACCGCGCTGTGCCGAGCATGGGACCTGCGCGG	2354		
DB	239	CCTCGACACCGGGCTGCCACCGCGCTGTGCCGAGCATGGGACCTGCGCGG	298		
QY	2355	AGTCGAGTGCAGCCCTGGCTGGATGGCGAACTGCACCATCGCTCACTATCTG	2414		
DB	299	AGTCGAGTGCAGCCCTGGCTGGATGGCGAACTGCACCATCGCTCACTATCTG	342		
QY	2415	GGGTAGTTAAAGGGTGGCTGGTGTGCAATGCGCAACGCGAGATGACCTTAG	2474		
DB	343	-----GNNN	391		
QY	2475	TGAATGTTGGCACTGGCTGCCAGCTGGGCTGGAGAGAGCTGGCTGTGACACT	2534		
DB	392	NN	451		
QY	2535	TGGAGACTGCTGCGGTGACAGCAAGACAAATGATGGAGATGGCTGGTGGATG	2594		
DB	452	NN	511		
QY	2595	ACCTGACTGCTGCTCCAGCCCTGTGCCATATCAACCGCTGTGGCTTGGCTCC	2654		
DB	512	ACCTGACTGCTGCTCCAGCCCTGTGCCATATCAACCGCTGTGGCTTGGCTCC	571		
QY	2655	ACCTCTGGACATCATCCAGGACACAGGTCCTGTGTGCACAGCAGAACCTAC	2714		
DB	572	ACCTCTGGACATCATCCAGGACACAGGTCCTGTGTGCACAGCAGAACCTAC	631		
QY	2715	TCTATGACCGGATCAAGTCTCTGCTGGGAGGACAGCAGACATAATCCCGG	2774		
DB	632	TCTATGACCGGATCAAGTCTCTGCTGGGAGGACAGCAGACATAATCCCGG	691		
QY	2775	ACCCCTTTGATGAGGGCATGCTGTGTTATTCGTGCGCAAGTGAATCAGATG	2834		
DB	692	ACCCCTTTGATGAGGGCATGCTGTGTTATTCGTGCGCAAGTGAATCAGATG	751		
QY	2835	CCCCCTGTTGTTGTAACATCAGTTTGTCAATAACCTCTCTTTGGATATACA	2894		
DB	752	CCCCCTGTTGTTGTAACATCAGTTTGTCAATAACCTCTCTTTGGATATACA	811		

QY	2895	GCAGCAAGATGGCAGCTTTGACTTTGGTGACAAATGGCGGCATCTCCATCATCT	2954		
DB	812	GCAGCAAGATGGCAGCTTTGACTTTGGTGACAAATGGCGGCATCTCCATCATCT	871		
QY	2955	TCGAGCGGCGACCTTTTCATCACAGGAGCACCGCTGTGGCTGCCATGGATCG	3014		
DB	872	TCGAGCGGCGACCTTTTCATCACAGGAGCACCGCTGTGGCTGCCATGGATCG	931		
QY	3015	TTGTCAATGGAACCATCATCATGAGACATGAGGAGATGAGATCCAGCTGTG	3074		
DB	932	TTGTCAATGGAACCATCATCATGAGACATGAGGAGATGAGATCCAGCTGTG	991		
QY	3075	GCAATTTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	3134		
DB	992	GCAATTTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1051		
QY	3135	CCTGTGACAGAGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	3194		
DB	1052	CCTGTGACAGAGAAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1111		
QY	3195	TCTGTGCTGCAAGATGAGGCTGAGCTACCTGACGACGCGGCGGCGGCGGCGG	3254		
DB	1112	TCTGTGCTGCAAGATGAGGCTGAGCTACCTGACGACGCGGCGGCGGCGGCGG	1171		
QY	3255	TCCTGAGGATCAGCTCACCCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	3314		
DB	1172	TCCTGAGGATCAGCTCACCCACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1231		
QY	3315	TGGTAGCGGCTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	3374		
DB	1232	TGGTAGCGGCTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1291		
QY	3375	ATTATTTCATTTGGGACAGACAGCTTACCAACGAGAGGCTTTGGGCTTCAGA	3434		
DB	1292	ATTATTTCATTTGGGACAGACAGCTTACCAACGAGAGGCTTTGGGCTTCAGA	1351		
QY	3435	CCTTTGTTCCGTGGGTTTGAATATGAATCTCCCGCAGATCTAATCTCTGGG	3494		
DB	1352	CCTTTGTTCCGTGGGTTTGAATATGAATCTCCCGCAGATCTAATCTCTGGG	1411		
QY	3495	GAAACAAGTCTGCGAGGCTATGAATTCAGCGCTCAAGCTTGGAGGATGGAG	3554		
DB	1412	GAAACAAGTCTGCGAGGCTATGAATTCAGCGCTCAAGCTTGGAGGATGGAG	1471		
QY	3555	ACAAACATCATGCCCTCAACATTCAGAGTGTATCTCTGCAAGAGGAAATGGG	3614		
DB	1472	ACAAACATCATGCCCTCAACATTCAGAGTGTATCTCTGCAAGAGGAAATGGG	1531		
QY	3615	AGTTTGTCTCAGCAGCCTCTCTCATTTGGGAGCATCATGGGCAATGGGCGG	3674		
DB	1532	AGTTTGTCTCAGCAGCCTCTCTCATTTGGGAGCATCATGGGCAATGGGCGG	1591		
QY	3675	GCATCTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	3734		
DB	1552	GCATCTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1651		
QY	3735	CCCTCACTGTGGCTCTGAGGGAGCCTCTATGTGGGTGATTTCAACTACAT	3794		
DB	1652	CCCTCACTGTGGCTCTGAGGGAGCCTCTATGTGGGTGATTTCAACTACAT	1711		
QY	3795	TCCTTCCCTCTGGAATGTCCAAACATCTTAGAGTGTAGGAAATAAGATTT	3854		
DB	1712	TCCTTCCCTCTGGAATGTCCAAACATCTTAGAGTGTAGGAAATAAGATTT	1750		
QY	3855	GTCAAGTCTGAGCACACAAATCTACCTGGCCACAGACCCCATGAGTGGGCGG	3914		
DB	1751	GTCAAGTCTGAGCACACAAATCTACCTGGCCACAGACCCCATGAGTGGGCGG	1810		
QY	3915	TTTCTGACAGCAACAGCGGCGGCTTTTAAATCAAGTCCACTGTGGTGAAG	3974		
DB	1811	TTTCTGACAGCAACAGCGGCGGCTTTTAAATCAAGTCCACTGTGGTGAAG	1870		
QY	3975	TTGTCAAGAACTCTGAGGTTGGTGGGCGGCGGCGGCGGCGGCGGCGGCGG	4034		

1871 TTTGTAAGAACTCTGAGGTGGTGGGGAGAGTGACAGTGCCTCCCTTTGATGACA 1930
4035 CTGCTGCGGGATGGTGGGAAGGCCACAGAAAGCCACACTCAACCAATCCAGGGGTATTA 4094
1931 CTGCTGCGGGATGGTGGGAAGGCCACAGAAAGCCACACTCAACCAATCCAGGGGCATTA 1990
4095 CAGTGGACAAGTTGGGTGATCTACTTCGTGGATGGGACCAATGATCAGAGCATCGATC 4154
1991 CAGTGGACAAGTTGGGTGATCTACTTCGTGGATGGGACCAATGATCAGAGCATCGATC 2050
4155 AGAATGGGATCATCTCCACCCCTGCTCGGCTCTAATGATCTCATCAGCCCGGCCACTCA 4214
2051 AGAATGGGATCATCTCCACCCCTGCTCGGCTCTAATGATCTCATCAGCCCGGCCACTCA 2110
4215 GCTGTGATCTGTCTGATGATATTTCCAGTGAAGTGGGACCAAGTGGGACCAAGTGG 4274
2111 GCTGTGATCTGTCTGATGATATTTCCAGTGAAGTGGGACCAAGTGGGACCAAGTGG 2170
4275 TCAACCAATGGACAACCTCACTTTATGCTCTGACAAACAAATGTGGTCTGCAAACTCTG 4334
2171 TCAACCAATGGACAACCTCACTTTATGCTCTGACAAACAAATGTGGTCTGCAAACTCTG 2230
4335 AAAACCAAGTGGCGATGTCGCGGGAGGCCCATGCACTGCGAGGTCCCTGGCATTTG 4394
2231 AAAACCAAGTGGCGATGTCGCGGGAGGCCCATGCACTGCGAGGTCCCTGGCATTTG 2290
4395 ACCACTCTCTGCTAAGCAAGTGGCCATCCAGCAACCTGAGTGCAGCCCGCTTTGG 4454
2291 ACCACTCTCTGCTAAGCAAGTGGCCATCCAGCAACCTGAGTGCAGCCCGCTTTGG 2350
4455 CTGTTTACACAAATGGGGTCTGTATATTGCTGAGACTGATGAGAAAGATCAACCGCA 4514
2351 CTGTTTACACAAATGGGGTCTGTATATTGCTGAGACTGATGAGAAAGATCAACCGCA 2410
4515 TCAGGAGGTCAACACTAGTGGAGATCTCACTGCTGCTGGGGCCCCAGTGGCTGTG 4574
2411 TCAGGAGGTCAACACTAGTGGAGATCTCACTGCTGCTGGGGCCCCAGTGGCTGTG 2470
4575 ACTGTAATAATGATGCCAATGATGATTTCTGGAGACATGATGTTATGCAAGGATG 4634
2471 ACTGTAATAATGATGCCAATGATGATTTCTGGAGACATGATGTTATGCAAGGATG 2530
4635 CAAAGTTAAATACCCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4694
2531 CAAAGTTAAATACCCCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2590
4695 ACCTTGGGACATCCGAATTCGTTTATCCGGAAGAACAGCCCTTCTCAACACCCAGA 4754
2591 ACCTTGGGACATCCGAATTCGTTTATCCGGAAGAACAGCCCTTCTCAACACCCAGA 2650
4755 ACATGTATGAGCTGTCTTCAACCAATTGACAGAGCTCTATCTGTTTGTATACACCGCA 4814
2651 ACATGTATGAGCTGTCTTCAACCAATTGACAGAGCTCTATCTGTTTGTATACACCGCA 2710
4815 AGCACTGTATACCCAAAGCTGCCACAGGAGTACCTGTACAACTTCACTACACTG 4874
2711 AGCACTGTATACCCAAAGCTGCCACAGGAGTACCTGTACAACTTCACTACACTG 2770
4875 GGGACGGGACATCACACTCATCACAGACAAATGGCAACATGGTAAGTGTCCGCGGAG 4934
2771 GGGACGGGACATCACACTCATCACAGACAAATGGCAACATGGTAAGTGTCCGCGGAG 2830
4935 ACTCTACTGGATGCCCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 4994
2831 ACTCTACTGGATGCCCTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2890
4995 TGGSCCAACAGTGCATCAAGAGTGTGACCAACAGGACACAGAGTGGCCATGATGA 5054
2891 TGGSCCAACAGTGCATCAAGAGTGTGACCAACAGGACACAGAGTGGCCATGATGA 2950
5055 CATACCATGGCAATTCGGCCTTCTGGCAACCAAAAGCAATGAAAACGGATGACAAAT 5114

2951 CATACCATGGCAATTCGGCCTTCTGGCAACCAAAAGCAATGAAAACGGATGACAAAT 3010
5115 TTTATGAGTACGACAGCTTTGGCGCTGACAAATGACCTTCCCTACTGGCCAGGTGA 5174
3011 TTTATGAGTACGACAGCTTTGGCGCTGACAAATGACCTTCCCTACTGGCCAGGTGA 3070
5175 GCAGTTCCTCCGAAGTGTATACAGACAGTTCAGTGTATGCTCCAGGTAGAGACCTCCAGCAAG 5234
3071 GCAGTTCCTCCGAAGTGTATACAGACAGTTCAGTGTATGCTCCAGGTAGAGACCTCCAGCAAG 3130
5235 ATGATGTACCAATACCAACCAACCTGTCTGCTCAGCGGCTTCTACACACTGCTGCAAG 5294
3131 ATGATGTACCAATACCAACCAACCTGTCTGCTCAGCGGCTTCTACACACTGCTGCAAG 3190
5295 ACCAAGTCCGGAACAGACTACTACATCGGGCCGATGCTCTTGGCGCTGCTGCTGCGCA 5354
3191 ACCAAGTCCGGAACAGACTACTACATCGGGCCGATGCTCTTGGCGCTGCTGCTGCGCA 3250
5355 ACGCATGGAGTGGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 5414
3251 ACGCATGGAGTGGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 3310
5415 CCGTGGGCAAGAGAAATGTCACGCTGCCATCGAACAGGCTCAACCTGGTGGAGTGGC 5474
3311 CCGTGGGCAAGAGAAATGTCACGCTGCCATCGAACAGGCTCAACCTGGTGGAGTGGC 3370
5475 GCCAGCGCAAGAGAGAGCTGCGGGCCAGGTCACTGTCTTGGCGCGGCTGCGGGTGC 5534
3371 GCCAGCGCAAGAGAGAGCTGCGGGCCAGGTCACTGTCTTGGCGCGGCTGCGGGTGC 3430
5535 ACAACCGAAATCTCTATCTCTGACCTTGTATCGGTAACACGACAGAGAAATCTATG 5594
3431 ACAACCGAAATCTCTATCTCTGACCTTGTATCGGTAACACGACAGAGAAATCTATG 3490
5595 ATGACCCAGCGCAAGTTCACCCCTCGGATTTCTGTACGACAGCGCGGGCGGCGGCTCT 5654
3491 ATGACCCAGCGCAAGTTCACCCCTCGGATTTCTGTACGACAGCGCGGGCGGCGGCTCT 3550
5655 GGTCAACAGCAGCAGGCTGAATGGTGTCAACGTGACATATCCCTCTGGGGGTTCATTG 5714
3551 GGTCAACAGCAGCAGGCTGAATGGTGTCAACGTGACATATCCCTCTGGGGGTTCATTG 3610
5715 CTGCATCCAGCGGGGATCATGTCTGAAAGAAATGGAATACGACAGCGGGCGGCGATCA 5774
3611 CTGCATCCAGCGGGGATCATGTCTGAAAGAAATGGAATACGACAGCGGGCGGCGATCA 3670
5775 CATCCAGGATCTTCGCTGATGGGAAGCATGGAGCTACACATCTTAGAGAAGCAGGTT 5828
3671 CATCCAGGATCTTCGCTGATGGGAAGCATGGAGCTACACATCTTAGAGAAGCAGGTT 3730
5829 CCATGGTCTCTACTACAGCCAGGAGCTATATCTTTGAGTTCGACAGAAATGACC 5888
3731 CCATGGTCTCTACTACAGCCAGGAGCTATATCTTTGAGTTCGACAGAAATGACC 3790
5889 GCCTCTCTCTGTGACAGTGCCTTCAACCTTCTACCTGGGCACTGGCGCGAGGCTGAT 5948
3791 GCCTCTCTCTGTGACAGTGCCTTCAACCTTCTACCTGGGCACTGGCGCGAGGCTGAT 3850
5949 TGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCTCAGTCAACAGGCT 6008
3851 TGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCTCAGTCAACAGGCT 3910
6009 TCACCTGAGGATGGGCACTCTTCAACCTTCTACCTGGGCACTGGCGCGAGGCTGAT 6068
3911 TCACCTGAGGATGGGCACTCTTCAACCTTCTACCTGGGCACTGGCGCGAGGCTGAT 3970
6069 ACAAGTATGGCAAACTGTCAAAGCTGGCAGAGAGCTCTATGACACCAACAGGCTGAT 6128
3971 ACAAGTATGGCAAACTGTCAAAGCTGGCAGAGAGCTCTATGACACCAACAGGCTGAT 4030
6129 TCACCTATGACGAGCGGCACTGCTGAGAACCAATCAACCTACAGAAATGAGGGCTTCA 6188
4031 TCACCTATGACGAGCGGCACTGCTGAGAACCAATCAACCTACAGAAATGAGGGCTTCA 4090

QY 6189 CCTGCACATCCGCTACCGTCAGATTGGGCGCCCTGATTGACCGACAGATCTTCGGCTTCA 6248
DB 4091 CCTGCACATCCGCTACCGTCAGATTGGGCGCCCTGATTGACCGACAGATCTTCGGCTTCA 4150
QY 6249 CTGAGGAAGCATGGTCAACGCCCGTTTGGACTACAACTATGACAAAGCTTCGGGGTGA 6308
DB 4151 CTGAGGAAGCATGGTCAACGCCCGTTTGGACTACAACTATGACAAAGCTTCGGGGTGA 4210
QY 6309 CCAGCATGCGAGGTGATCAACGAGACCCCATGCCCATTTGATCTCTATCGCTATGATG 6368
DB 4211 CCAGCATGCGAGGTGATCAACGAGACCCCATGCCCATTTGATCTCTATCGCTATGATG 4270
QY 6369 ATGTGTGAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCATTTACTATGACATTAAAC 6428
DB 4271 ATGTGTGAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCATTTACTATGACATTAAAC 4330
QY 6429 AGATCATCACACAGCTGTATGACCCACACCAAGCATTTTGTATGATATGCGAGATGA 6488
DB 4331 AGATCATCACACAGCTGTATGACCCACACCAAGCATTTTGTATGATATGCGAGATGA 4390
QY 6489 AGAAGTGCAGTATGAGATCTTCGGCTCGCTCATGTACTGGATGACCGTCCAGTATGATA 6548
DB 4391 AGAAGTGCAGTATGAGATCTTCGGCTCGCTCATGTACTGGATGACCGTCCAGTATGATA 4450
QY 6549 ACATGGGCGAGTAGTGAAGAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCT 6608
DB 4451 ACATGGGCGAGTAGTGAAGAGGAGCTGAAGGTAGGACCCCTACGCCAATACCACTCGCT 4510
QY 6609 ACTCCATGATGATGATGTCGCGGCGAGCTGCGAGACAGTCTCCATCAATGACAGCCAC 6668
DB 4511 ACTCCATGATGATGATGTCGCGGCGAGCTGCGAGACAGTCTCCATCAATGACAGCCAC 4570
QY 6669 TCTGGGCGTACAGCTACGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTG 6728
DB 4571 TCTGGGCGTACAGCTACGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGGGAACAGTG 4630
QY 6729 CAGGGCTCACACCACTACGATATGATATCGCGACCGCATCACTCGGCTGGGTGACGTGC 6788
DB 4631 CAGGGCTCACACCACTACGATATGATATCGCGACCGCATCACTCGGCTGGGTGACGTGC 4690
QY 6789 AATACAAGATGAGTATGAGTGGCTTCTGAGGACGCGGCGGTGATATCTTTGATGACA 6848
DB 4691 AATACAAGATGAGTATGAGTGGCTTCTGAGGACGCGGCGGTGATATCTTTGATGACA 4750
QY 6849 ACTCAGTGGCTGCTCATCAAGCCCTACAACCGGCTGGCAGCTGGAGTGTCAAGGTACC 6908
DB 4751 ACTCAGTGGCTGCTCATCAAGCCCTACAACCGGCTGGCAGCTGGAGTGTCAAGGTACC 4810
QY 6909 GCTACAGTGGCTGGGCGGCGGTCTCCAGCAAGACAGCCACAGCCACCACTTGCAGT 6968
DB 4811 GCTACAGTGGCTGGGCGGCGGTCTCCAGCAAGACAGCCACAGCCACCACTTGCAGT 4870
QY 6969 TCTTCTATGACAGCTGACCAACCCCAAGGTGACCCACCTGTACAAACCACTCCAGT 7028
DB 4871 TCTTCTATGACAGCTGACCAACCCCAAGGTGACCCACCTGTACAAACCACTCCAGT 4930
QY 7029 CTGAGATCACTCCCTCTACTAGCTTGCAGGACCACTCTTTGGCATGGAGCTGAGGA 7088
DB 4931 CTGAGATCACTCCCTCTACTAGCTTGCAGGACCACTCTTTGGCATGGAGCTGAGGA 4990
QY 7089 GTGGTATGAGTTTATACATAGTTTGTGACAAATCGGAGCCCTCTTGTGCTTTTAGTG 7148
DB 4991 GTGGTATGAGTTTATACATAGTTTGTGACAAATCGGAGCCCTCTTGTGCTTTTAGTG 5050
QY 7149 GAACAGTTTATGATCAAGCAAAATCTCTGTACAGCCCTATGGGAGATCTACATGGATA 7208
DB 5051 GAACAGTTTATGATCAAGCAAAATCTCTGTACAGCCCTATGGGAGATCTACATGGATA 5110
QY 7209 CCAACCCCACTTTACATATCATATAGGCTACCATGGTGGCTCTATGATCCACTCACCA 7268
DB 5111 CCAACCCCACTTTACATATCATATAGGCTACCATGGTGGCTCTATGATCCACTCACCA 5170

QY 7269 AGCTTGTCACATGGCGCGGAGATTATGATGTGCTGGCGGACGCTGGACTAGCCAG 7328
DB 5171 AGCTTGTCACATGGCGCGGAGATTATGATGTGCTGGCGGACGCTGGACTAGCCAG 5230
QY 7329 ACCAGAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCCCTTTAACTCTCTATATGT 7388
DB 5231 ACCAGAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCCCTTTAACTCTCTATATGT 5290
QY 7389 TCAAAAACAAACCCCATCAGCAACTCCACAGGACATCAAGTGCTTCATACAGATGTTA 7448
DB 5291 TCAAAAACAAACCCCATCAGCAACTCCACAGGACATCAAGTGCTTCATACAGATGTTA 5350
QY 7449 ACAGCTGGCTCTCACCTTTGGATTCCAGCTTACAAACGTCATCCCTGGTTATCCCAAC 7508
DB 5351 ACAGCTGGCTCTCACCTTTGGATTCCAGCTTACAAACGTCATCCCTGGTTATCCCAAC 5410
QY 7509 CAGACATGATGCCATGGAAACCTCTTACGAGCTCATCCACACAGATGAAACGCGAG 7568
DB 5411 CAGACATGATGCCATGGAAACCTCTTACGAGCTCATCCACACAGATGAAACGCGAG 5470
QY 7569 AGTGGGACACAGCAAGTCTATCTCGGGGTACAGTGTGAAGTACAGAACGACGCTCAAG 7628
DB 5471 AGTGGGACACAGCAAGTCTATCTCGGGGTACAGTGTGAAGTACAGAACGACGCTCAAG 5530
QY 7629 CCTTTGTCACTTATAGAACGGTTTGACAGCTCTATGGCTCCACAATCACCAGCTGCCAG 7688
DB 5531 CCTTTGTCACTTATAGAACGGTTTGACAGCTCTATGGCTCCACAATCACCAGCTGCCAG 5590
QY 7689 AGGCTCCAAAGACCAAGAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGT 7748
DB 5591 AGGCTCCAAAGACCAAGAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGT 5650
QY 7749 TTGCTTTGAAGGATGGCGGAGTGACACACACATCATCAGTGTGGCCCAATGAGGATGGGC 7808
DB 5651 TTGCTTTGAAGGATGGCGGAGTGACACACACATCATCAGTGTGGCCCAATGAGGATGGGC 5710
QY 7809 GAAGGTTGTGCCATCTTGAAACATGCCCACTACCTAGAGAACCTGCACCTTCCACATTTG 7868
DB 5711 GAAGGTTGTGCCATCTTGAAACATGCCCACTACCTAGAGAACCTGCACCTTCCACATTTG 5770
QY 7869 ATGGGTTGGATACCATTTACTTTGTGAACACAGACCTTCAGAAAGTGCCTCGCCATCC 7928
DB 5771 ATGGGTTGGATACCATTTACTTTGTGAACACAGACCTTCAGAAAGTGCCTCGCCATCC 5830
QY 7929 TGGGCTCAGTGGGGGCGCGAACCTTGAGAAATGGGTCAAACGTCATGTGTCCAGA 7988
DB 5831 TGGGCTCAGTGGGGGCGCGAACCTTGAGAAATGGGTCAAACGTCATGTGTCCAGA 5890
QY 7989 TCACACAGTACTTATGGCAGACTAGACGCTTACACAGACATCCAGCTCCAGTACGGGG 8048
DB 5891 TCACACAGTACTTATGGCAGACTAGACGCTTACACAGACATCCAGCTCCAGTACGGGG 5950
QY 8049 CACTGTCTTTGAACACACACGCTACGGGACAACTTGGATGAGGAGAGGACCGGCTCTCG 8108
DB 5951 CACTGTCTTTGAACACACACGCTACGGGACAACTTGGATGAGGAGAGGACCGGCTCTCG 6010
QY 8109 AGTGGGCGGCGAGAGAGCGTGGCGCAAGCTGGGCGCGGAGAGAGAGACTCGGG 8168
DB 6011 AGTGGGCGGCGAGAGAGCGTGGCGCAAGCTGGGCGCGGAGAGAGAGACTCGGG 6070
QY 8169 AAGGGAGGAGGCGCTGCGGCGCTGGACAGAGGGGAGAGAGAGAGTCTGAGACAG 8228
DB 6071 AAGGGAGGAGGCGCTGCGGCGCTGGACAGAGGGGAGAGAGAGAGTCTGAGACAG 6130
QY 8229 GCGGGTGCAGGCTACGAGCGCTTTTGTGATCTCTGTGAGAGAGTACCCAGAACTGT 8288
DB 6131 GCGGGTGCAGGCTACGAGCGCTTTTGTGATCTCTGTGAGAGTACCCAGAACTGT 6190
QY 8289 CAGACAGCGCAACCAACATCCACTTATGAGACAGAGCGAGATGGGCGGAGGTGA 8344
DB 6191 CAGACAGCGCAACCAACATCCACTTATGAGACAGAGCGAGATGGGCGGAGGTGA 6246

Fri Jun 25 08:41:51 2004

RESULT 2	AY413476	5970 bp	DNA	linear	GSS 12-DEC-2003
LOCUS	AY413476				
DEFINITION	Pan troglodytes HMC4903 gene, VIRTUAL TRANSCRIPT, partial sequence.				
ACCESSION	AY413476				
VERSION	AY413476.1	GI:39769438			
KEYWORDS	GSS.				
SOURCE	Pan troglodytes (chimpanzee)				
ORGANISM	Pan troglodytes				
REFERENCE	Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pan.				
AUTHORS	1 (bases 1 to 5970)				
TITLE	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
JOURNAL	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios				
PUBLISHED	Science 302 (5652), 1960-1963 (2003)				
REFERENCE	14671302				
AUTHORS	2. (bases 1 to 5970)				
TITLE	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreria,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
JOURNAL	Direct Submission				
COMMENT	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
FEATURES	This sequence was made by sequencing genomic exons and ordering them based on alignment.				
source	Location/Qualifiers				
gene	1..5970				
ORIGIN	/organism="Pan troglodytes"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9598"				
	<1..>5970				
	/locus_tag="HCM4903"				
Query Match	60.4%; Score 5045.8; DB 29; Length 5970;				
Best Local Similarity	85.6%; Pred. No. 0;				
Matches 5149; Conservative	0; Mismatches 815; Indels 54; Gaps 3;				
Qy	2333 CATGGGACCTGCCGCGACGGCAAGTCGAGTCAGCCCTGGCTGGGAATCGCAACACTGC 2392				
Db	1 CATGGGACCTGCCGCGACGGCAAGTCGAGTCAGTCCTGGCTGGGAATCGCAACACTGC 60				
Qy	2393 ACCATCGCTCACTATCTGGATAGGTAGTTAAAGAGGGTTGCCTGGTGTGCAATGCG 2452				
Db	61 ACCATCGNN 112				
Qy	2453 AACGGCAGATGTACCTTAGACCTGAATGTTGGCAGTCGCTCTGCCAGCTGGCGTGGAGA 2512				
Db	113 -----NN 153				
Qy	2513 GGAGCTGGCTGTGACACTTCATCGAGACTGCTCGGTGACAGAAACAATGATGGA 2572				
Db	154 NNN 213				
Qy	2573 GATGGCTGTGGAGTGCATCGAGCCCTGACTGTGCTCCAGCCCTGTCGCATATCAAC 2632				
Db	214 NATGGCTGTGGAGTGCATCGAGCCCTGACTGTGCTCCAGCCCTGTCGCATATCAAC 273				
Qy	2633 CCGCTGTGCCCTGGCTCCCTTAACCCCTCTCGAATCATCCAGGAGACACAGGTCCTGTG 2692				
Db	274 CCGCTGTGCCCTGGCTCCCTTAACCCCTCTCGAATCATCCAGGAGACACAGGTCCTGTG 333				
Qy	2693 TCACAGCAGAACCTACACTCTTCTATGACCGCATCAAGTTCCTCGTGGCGACGGACAGC 2752				
Db	334 TCACAGCAGAACCTACACTCTTCTATGACCGCATCAAGTTCCTCGTGGCGACGGACAGC 393				
Qy	2753 ACGCACATAATCCCGGGGAGAACCCCTTTGATGGAGGGATGCTGTGTATTTCGTGGC 2812				

QY 6047 GGCACTGGCGCAGGGTGATATCAAGTATGGCAAACTGTCAAAGCTGGCGAGAGCGCTC 6106
Db |||||
QY 3673 GGCACTGGCGCAGGGTGATATCAAGTATGGCAAACTGTCAAAGCTGGCGAGAGCGCTC 3732
Db |||||
QY 6107 TATGACACCAACCAGGTCAGTTTACCTATGACGAGACGGCAGGCATGCTGGAAGACCATC 6166
Db |||||
QY 3733 TATGACACCAACCAGGTCAGTTTACCTATGACGAGACGGCAGGCATGCTGGAAGACCATC 3792
Db |||||
QY 6167 AACCTACAGATGAGGGCTTACCTGACCATCGCTACCGTACAGATGGGCGCCCTGATT 6226
Db |||||
QY 3793 AACCTACAGATGAGGGCTTACCTGACCATCGCTACCGTACAGATGGGCGCCCTGATT 3852
Db |||||
QY 6227 GACCGACAGATCTTCCGCTTCACTGAGGAAGGATGCTCAACGCCCGTTTGTACTACAAC 6286
Db |||||
QY 3853 GACCGACAGATCTTCCGCTTCAACCGAGGAAGGCATGCTCAACGCCCGTTTGTACTACAAC 3912
Db |||||
QY 6287 TATGACAAACAGCTTCCGGGTGACAGCATGACAGCTGTGATCAACGAGACCCCACTGCC 6346
Db |||||
QY 3913 TATGACAAACAGCTTCCGGGTGACAGCATGACAGCTGTGATCAACGAGACCCCACTGCC 3972
Db |||||
QY 6347 ATTGATCTCTATCGCTATGATGATGCTCAGGCAAGACAGAGCAGTTTGGGAGTTTGGT 6406
Db |||||
QY 3973 ATTGATCTCTATCGCTATGATGATGCTCAGGCAAGACAGAGCAGTTTGGGAGTTTGGT 4032
Db |||||
QY 6407 GTCAATTATGATGACATTAACAGATCATCAACACAGCTGTCTATGACCCACACCAAGCAT 6466
Db |||||
QY 4033 GTCAATTATGATGACATTAACAGATCATCAACACAGCTGTCTATGACCCACACCAAGCAT 4092
Db |||||
QY 6467 TTTGATGCAATGGCAGGATGAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTAC 6526
Db |||||
QY 4093 TTTGATGCAATGGCAGGATGAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTAC 4152
Db |||||
QY 6527 TGGATGACCCCTCCAGTATGATTAACATGGGGCGAGTAGTGAAGAGGAGCTGAAGGTAGGA 6586
Db |||||
QY 4153 TGGATGACTGCCAGTATGATTAACATGGGGCGAGTAGTGAAGAGGAGCTGAAGGTAGGA 4212
Db |||||
QY 5587 CCCTAGCCCAATACCACTCCTCTATGATGATGATGCTGACGGCAGCTGACAGACA 6646
Db |||||
QY 4213 CCCTAGCCCAATACCACTCCTCTATGATGATGATGCTGACGGCAGCTGACAGACA 4272
Db |||||
QY 6647 GTCTCCATCAATGACAAAGCCACTCTGCGCTACAGCTACAGCTCAATGGGACCTGAC 6706
Db |||||
QY 4273 GTCTCCATCAATGACAAAGCCACTCTGCGCTACAGCTACAGCTCAATGGGACCTGAC 4332
Db |||||
QY 6707 TTACTGAGCCCTGGGAACAGTGCACGCTCACACCCTACGGTATGACATCCGCGACCCG 6766
Db |||||
QY 4333 TTACTGAGCCCTGGGAACAGTGCACGCTCACACCCTACGGTATGACATCCGCGACCCG 4392
Db |||||
QY 6767 ATCACTCGGCTGGGTGACGTGCAATACAGATGGATGAGGATGGCTTCTGAGGACGG 6826
Db |||||
QY 4393 ATCACTCGGCTGGGTGACGTGCAATACAGATGGATGAGGATGGCTTNNNTGAGGACGG 4452
Db |||||
QY 6827 GCGGCTGATATCTTTGAGTACAACCTCAGCTGGCTGCTCATCAAGGCTTACAAACCGGCT 6886
Db |||||
QY 4453 GCGGCTGATATCTTTGAGTACAACCTCANCTGNNNTANNATCAAGGNCAACCGGCT 4512
Db |||||
QY 5887 GCGAGCTGAGGTGACGTTACCGCTACGATGGCGGCGCGCTGTCCAGCAAGAGC 6946
Db |||||
QY 4513 GCGAGCTGAGGTGACGTTACCGCTACGATNNNNNTGGGCGGCGGTGTCCAGCAAGAGC 4572
Db |||||
QY 6947 AGCCACAGGCCACCACTGCACTTCTTATGACAGCTTGACCAACCCACCHAGTCAAC 7006
Db |||||
QY 4573 AGCCACAGGCCACNN 4632
Db |||||
QY 7007 CACTGTGTAACCACTCCAGCTTGAGATCACTCCCTCTACTACGACTTGAAGGACAC 7066
Db |||||
QY 4633 NNN 4692
Db |||||
QY 7067 CTCTTTGCCATGAGCTGACGAGTGTGATGAGTTTATAGCTTTGACAAACATCGGG 7126
Db |||||
QY 4693 CTCTTTGCCATGAGCTGACGAGTGTGATGAGTTTATAGCTTTGACAAACATCGGG 4752
Db |||||
QY 7127 ACCCCTCTCTCTCTTTAGTGAACACAGGTTTGTGATCAAGCAAACTCTGTACACAGCC 7186
Db |||||

Db 4753 ACCCCTCTCTCTCTTTAGTGAACACAGGTTTGTGATCAAGCAAACTCTGTACACAGCC 4812
QY |||||
QY 7187 TATGGGAGATCTCATGATACCAACCCCACTTTTCAAGATCATATAGGCTACCATGGT 7246
Db |||||
QY 4813 TATGGGAGATCTCATGATACCAACCCCACTTTTCAAGATCATATAGGCTACCATGGT 4872
Db |||||
QY 7247 GGCCTCTATGATCTCACTCAACAGCTTGTCCACATGGGCGGCGAGATTTATGATGCTG 7306
Db |||||
QY 4873 GGCCTCTATGATCTCACTCAACAGCTTGTCCACATGGGCGGCGAGATTTATGATGCTG 4932
Db |||||
QY 7307 GCGGACCGCTGGACTAGCCAGACCAACAGCTGTGGAAGACCTTTAGTAGCAGCAACCTC 7366
Db |||||
QY 4933 GCGGACCGCTGGACTAGCCAGACCAACAGCTGTGGAAGACCTTTAGTAGCAGCAACCTC 4992
Db |||||
QY 7367 ATGCTCTTTAATCTCTATGTTCAAAAACCAACCCCATCAGCAACTCCAGGACATC 7426
Db |||||
QY 4993 ATGCTCTTTAATCTCTATGTTCAAAAACCAACCCCATCAGCAACTCCAGGACATC 5052
Db |||||
QY 7427 AAGTCTTTCAAGACAGATGTTAAACAGCTGGCTGCTCACTTTTGGATTTCCAGCTACACAC 7486
Db |||||
QY 5053 AAGTCTTTCAAGACAGATGTTAAACAGCTGGCTGCTCACTTTTGGATTTCCAGCTACACAC 5112
Db |||||
QY 7487 STGATCCCTGGTTATCCCAACCAAGACATGGATGCCATGGAACCCCTCTACGAGCTATC 7546
Db |||||
QY 5113 STGATCCCTGGTTATCCCAACCAAGACATGGATGCCATGGAACCCCTCTACGAGCTATC 5172
Db |||||
QY 7547 CACACACAGATGAACACCGAGGTGGGACCAACAGCAAGTCTATCTCCGGGTACAGTGT 7606
Db |||||
QY 5173 CACACACAGATGAACACCGAGGTGGGACCAACAGCAAGTCTATCTTTGGGGTACAGTGT 5232
Db |||||
QY 7607 GAAGTACAGAGCAGCTCAAGGCTTTGTCACTTTAGAACGGTTTGACAGCTCTATGGC 7666
Db |||||
QY 5233 GAAGTACAGAGCAGCTCAAGGCTTTGTCACTTTAGAACGGTTTGACAGCTCTATGGC 5292
Db |||||
QY 7667 TCCCAATATCAACAGCTGCCAGAGCTCCAAAGACCAAGAGTTTGCATCCAGCGGCTCA 7726
Db |||||
QY 5293 TCCCAATATCAACAGCTGCCAGAGCTCCAAAGACCAAGAGTTTGCATCCAGCAGCTCA 5352
Db |||||
QY 7727 GTCTTTGGCAAGGGGTCAAGTTTGCCTTTGAAGGATGSCCGAGTACCAACAGACATCATC 7786
Db |||||
QY 5353 GTCTTTGGCAAGGGGTCAAGTTTGCCTTTGAAGGATGSCCGAGTACCAACAGACATCATC 5412
Db |||||
QY 7787 AGTGTGSCCAATAGGATGGGCGAAGGTTGCTGCCATCTTTGAACCACTGCCCACTACCTA 7846
Db |||||
QY 5413 AGTGTGSCCAATAGGATGGGCGAAGGTTGCTGCCATCTTTGAACCACTGCCCACTANCA 5472
Db |||||
QY 7847 GAGAACCTGCATCTTCACTTTGATGGGTGATACCCATTTACTTTTGAACACCGAGCT 7906
Db |||||
QY 5473 GAGAACCTGCATCTTCACTTTGATGGGTGATACCCATTTACTTTTGTGAACCCAGGACCT 5532
Db |||||
QY 7907 TCAGAAAGGTGACCTGGCCATCTGGGCTCAGTGGGGGGCGGAAACCCCTGGAGAAATGGG 7966
Db |||||
QY 5533 TCAGAAAGGTGACCTGGCCATCTGGGCTCAGTGGGGGGCGGAAACCCCTGGAGAAATGGG 5592
Db |||||
QY 7967 GTCAACGCTACTGTGTCCAGATCAACAGATCTTAATGGCAGGACTAGACGCTACACA 8026
Db |||||
QY 5593 GTCAACGCTACTGTGTCCAGATCAACAGATCTTAATGGCAGGACTAGACGCTACACA 5652
Db |||||
QY 8027 GACATCCAGCTCCAGTACGGGCACTGTGCTTGAACACACGCTACGGGACAACTGTGAT 8086
Db |||||
QY 5653 GACATCCAGCTCCAGTACGGGCACTGTGCTTGAACACACGCTACGGGACAACTGTGAT 5712
Db |||||
QY 8087 GAGGAGAAAGGCAACCGGCTCTGGAGTGGCCCGGAGAGAGCCGTGCGCAAGCGTGGGCC 8146
Db |||||
QY 5713 GAGGAGAAAGGCAACCGGCTCTGGAGTGGCCCGGAGAGAGCCGTGCGCAAGCATGGGCC 5772
Db |||||
QY 8147 GCGGAGCAGAGAGACTGCGGAAAGGGGAGAAAGCCCTCGGGCTGGACAGAGGGGAG 8206
Db |||||
QY 5773 GCGGAGCAGAGAGACTGCGGAAANNAGAGAGCCCTGCGGGCTGGACAGAGGGGAG 5832
Db |||||
QY 8207 AAGCAGAGGCTGCTGAGCAGACAGGGCGGTGCAAGGCTACAGCGCTTTTCTGTATCTCT 8266
Db |||||

Db 5833 AAGCAGCAGGTGTGAGCAGACAGGCGGGTGAAGGCTACGACGGCTTTTTCGTGATCTCT 5892

Qy 8267 GTCCAGCAGTACCCAGAACCTGTGAGCAGCGGCCAACAACTTCCATCTTCATGAGACAGAGC 8326

Db 5893 GTCCAGCAGTACCCAGAACCTGTGAGCAGCGGCCAACAACTTCCATCTTCATGAGACAGAGC 5952

Qy 8327 GAGATGGCGCGAGGTGA 8344

Db 5953 GAGATGGCGCGAGGTGA 5970

RESULT 3

AY405420 5094 bp DNA linear GSS 12-DEC-2003

LOCUS

DEFINITION Homo sapiens HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.

ACCESSION AY405420

VERSION AY405420.1 GI:39761394

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 5094)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Givello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trics

JOURNAL Science 302 (5652), 1960-1963 (2003)

PubMed 14671302

REFERENCE 2 (bases 1 to 5094)

Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Givello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source 1..5094

locus_tag="HCM2218"

gene

ORIGIN

Query Match 28.28; Score 2353; DB 29; Length 5094;

Best Local Similarity 67.04; Pred. No. 0;

Matches 3442; Conservative 0; Mismatches 1647; Indels 45; Gaps 6;

Qy 2912 TTTGACTTGTGTGACAAATGGCGGCATCTCCATCATCTCGGTTTCAGCGGGCACCTTTTC 2971

Db 2 TTTGACTTGTGTGACAAATGGCGGCCTCTTAACCTTGTGATTTGAACGATCCCATTC 61

Qy 2972 ATCACAGAGACACACCCCTGTGGTCCCATGGGATCGCTTTCTTGTGATGGAACCATC 3031

Db 62 CTCACCTCAGTATCATCTGTGTGGATTCATGGAATGTCTTTATGTGATGATACCCCTA 121

Qy 3032 ATCATGAGATCAGGAGATGAGATTCACGCTGTGACCTGAGCAATTTTGGCCGCC 3091

Db 122 GTCATGAGAAAGAGAGATGACATTCACGCTGTGATCTGAGTGGATTCGTGAGGCCA 181

Qy 3092 AACCCAGTCTGTCTCATCCACATGACGTCTTCCGACGTCTCTGTGACAGAAAGGC 3151

Db 182 AATCCCATCATCTGTGTGATCATCACCTTTATCCACCTTTTTCAGATCTTCTCCTGAAGACAGT 241

Qy 3152 CCATTGTGCGGAAATTCAGGCTTTGCGAGGAGAAATCTCTATCTCTGCTGCAAGATG 3211

242 CCCATCATTCGCCGAAACACANGTACTCCAGAGGAAACTCAATTCAGGAAACAGATTG 301

3212 AGGCTGAGCTACCTGAGCAGCGGACCCCTGGCTTACAAATCTGCTCGAGGATCAGCCTC 3271

302 AAATCTCTCTACTTACTGTTCCAGAGCTGCGAGGTATTAAGTCAGTTCTCAGATCACCATTG 361

3272 ACCCACCGACCATCCCTTCAACCTCATGAAGTGCACCTCATGGTAGCGGTGGAGGC 3331

362 ACCCAGTCTATTATTCCATTTAATTAATGAAGTTCATCTTATGCTAGCTAGTAGGA 421

3332 CGCCTCTTCAGGAAGTGTTCGCTGCGAGCCCGACCTGCTCTATTATTTCATTTGGGAC 3391

422 AGATCTTTTCAAAAGTGGTTTCTGCTCTACAAACTTGGCCCTATCTTTCATATGGGAT 481

3392 AAGACAGACGCTTCAACACAGAGGTGTTTGGGCTTTTCAGAAAGCCTTTGTTCCGTGGGT 3451

482 AAAACAGATGCATATAATCAGAAAGTCTATGGTCTATCTGAAGCTGTGTGTCAGTTGA 541

3452 TATGAATATGAATCCTGCCAGATCTAATCTGTGGGAAAAAGAACAAACAGTCTGCAG 3511

542 TATGAGTATGATGCTGTTTGGACCTGACTCTGTGGGAAAAAGAGGACTGCCATTTGCGAG 601

3512 GGCTATGAAATTTGACGCGTCCAAGCTTGGAGGATGGAGCCTAGACAAACATCATGCCCTC 3571

602 GGCTATGAATTTGGATGGTCCACATCGGTGGTGGACATTAGATAACATCAGTGTCTG 661

3572 AACATTCAAAGTGTATCTTCGACAAAGGAAATGGGGAGAACCAAGTTTGTGTCTCAGCAG 3631

662 GATGTACAGAACGCTATCTGTACAGGAAACCGGGGAAACCCAGTTTCATCTCCACGAG 721

3632 CCTCTGTCTATTTGGAGCATCATGGGCAATGGCGCGGAGAGAGCATCTCTCTGCCCGAGC 3691

722 CTCTCAGTGTGATGATCATCTGGGCAATGGCGGAGGCGCAGCATTTCTGCCCCAGT 781

3692 TGCAACGCGCTTGTGACGCGCAACAAAGCTCTCTGGCCCGACCTGCGCTCAGCTGTGGTCT 3751

782 TGCAATGGTCAAGCTGATGGTAAACAAGTTACTTGGCCCCAGTGGCGCTAGCTTTGGGATC 841

3752 GACGGAGCCTCTATTTGGGTGATTTCAACTACATTAGAAGGATCTTCCCTCTGGAAT 3811

842 GATGGAGTGTGATGAGTAGGAGATTTCAACTATGTGCGCGGATTTTCCCTCTTGGAAAT 901

3812 GTCAACCAACATCTTAGAGCTGAGGAATAAAGATTTTCAGACATAGTCACAGTCAGCACAC 3871

902 GTAACAGTGTCTTAGAAGTAAAGCA-----GCAACCCAGCTCAT 940

3872 AAATATCTGCGGCCACAGACCCCATAGTGGGGCCCTCTTCTTCTGACAGCAACAGC 3931

941 AGATACTACTTTGCAACGGATCCAGTCAAGGAGATCTGACGTTTCTGACAAACACC 1000

3932 CGGCGGGTCTTTAAATCAAGTCCACTGTGTGGTGAAGACCTTGTCAAGAACTCTGAG 3991

1001 CGCAGATTTATCCCCAAAGTCACTTACGGGGCCAAAGACTTGACTAAATGCGAGAA 1060

3992 GTGTTGCGGGGACAGGTGACAGTGTCTCTTCTTGTGATGACATCTCTGCGGGGATGGT 4051

1061 GTCTGCGAGGACAGGGGAGCAATGCTTCCGTTTACGAGGCGAGATGTGGGGATGGA 1120

4052 GGGAGGCCACAGAAAGCCACACTCAACATCCCAAGGGGTATTACAGTGGACAAGTTTGGG 4111

1121 GGGAGGCGGTGGAAAGCCACTCATGAGTCCCAAGAGATGGCAGTTGATAGAATGGA 1180

4112 CTGATCTACTTCTGTGATGGCACCATGATCAGAGCGATCGATCAGATGGATCATCTCC 4171

1181 TTAATCTACTTTTGTGATGGAACCATGATTAGGAAGTTGACCAAAATGAAATCATATCA 1240

4172 ACCCTGTCTGCTCTATGATCTCAGATCAGCCCGGCACTCAGCTGTGATTTGTCTCATG 4231

1241 ACTCTTCTGGGCTCTAAGATTTGACTTTCAGCCAGACCTTTTAACTTGTGACACCGAGCATG 1300

4232 GATATTTTCCAGGTAAGACTGGAGTGGCCGCCACAGACTTAGCCATCAACCCAAATGGACAAC 4291

1301 CACATCAGCCAGGTACGCTTGGAAATGCCCCACTGACCTAGCCATTAAACCTTATGGATAAC 1360

Db 3518 ATGACCTATACGAGGACCTTTGATGCTCATGGCCGATCAAGGAGATTCAATATGAGATA 3577
 QY 6509 TTCGGCTCGCTCATGTACTGATGAGACCGTCCTAGTATGATAAATGAGGGCGAGTAGTGAAG 6568
 Db 3578 TTCAGGTGCGTCTATGTACTGATGATCAAAATTCAGTATGATAAATGAGGGTGGTAAACCAAG 3637
 QY 6569 AAGGAGCTGAAGTGAAGACCTAGCCCAATACCACTCGTACTCTCTATGAGTATGATGCT 6628
 Db 3638 AGAGAGATTAAATAGGGCCCTTTGGCAACACCAAAATATGCTTATGAATATGATGTT 3697
 QY 6629 GAGGGCAGCTGAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGAC 6688
 Db 3698 GATGGACAGCTCCAAACAGTTTACCTCAATGAAAGATAATGTGGGGGTACAACTACGAT 3757
 QY 6689 CTCAATGGGAACCTGACACTTACTGAGCCCTGGGAACAGTGCAGCGCTCACACCACTACGG 6748
 Db 3758 CTGAATGGAAACCTCCATTTACTGAAACCCCAAGTAACAGTGCAGCGCTGACACCCCTTCGC 3817
 QY 6749 TATGACATCCGCAACCGCATCACTCGGCTGGGTGAGCTGCAATACAAAGATGATGAGGAT 6808
 Db 3818 TATGACCTGCGAGACAGATCACTCGACTGGGTGATGTTCAATATCGGTGGATGAAGAT 3877
 QY 6809 GGCTTCCTGAGGAGCGGGCGGTGTATATCTTTGAGTACAATCTGAGCTGGCTGCTCAATC 6868
 Db 3878 GGTTCCTACGTCAAAAGGGCGCGAAATCTTTGAAATATAGCTCCAAAGGGGCTTTAACT 3937
 QY 6869 AAGGCTTACAAACCGGCTGCGACCTGGAGTGTGAGTACCGCTACGATGGCTGGGGGG 6928
 Db 3938 CGAGTTTACAGTAAAGGAGTGGCTGAGCAGTGTCTACCGTTATGACGCGCTTGGGAAG 3997
 QY 6929 CGGCTGTCCAGAGAGACCCACAGCCACCTGAGTCTCTCTATGACAGCTGACCC 6988
 Db 3998 CGTGTCTTCTAGCAAAACCACTAGGACAGCACCTGCAGTTTTTTATGCTGACTTAAT 4057
 QY 6989 AACCCCAAGAGTCAACCCACCTGACACCACTCCAGCTCTGAGATCACTCCCTCTAC 7048
 Db 4058 TATCCCACTAGGATTAATCTATGCTCAACCACTTCAGTTTCAGAAATTAACCTCCGTAT 4117
 QY 7049 TAGGACTTGAAGAGACACCTCTTTGGCATGAGCTGAGCAGTGGTGTATGAGTTTACATA 7108
 Db 4118 TATGATCTCCAGGACATCTTTTGGCATGGAATCAGCAGTGGGATGAATCTATATT 4177
 QY 7109 GCTTGTGACAACTTCGGGACCCCTCTTGTGCTCTTTAGTGGACAGGTTTGTATGATCAAG 7168
 Db 4178 GCATCGGATAACAGGAGCACCACTGGCTGTGTTTCAAGTAAATGGGCTTATGCTGAA 4237
 QY 7169 CAAATCTGTACACAGCCTATGGGAGATCTACATGGATACCAACCCCACTTTCAGATC 7228
 Db 4238 CAGATTGACGACTGATGATGGGAAATCTATTTTGTCTCTATGATGACTTTCAGCTG 4297
 QY 7229 ATCATAGGCTACCATGGTGGCTCTATGATGCTACTCACCAGCTTGTCCATGAGGCGCG 7288
 Db 4298 GTAAATGGATTTCATGGTGGCTGTATGACCACTACCAAAATTAATCCACTTTGGAGAA 4357
 QY 7289 CGAGATTATGATGCTGGCGGACCTGGATGAGCCAGACACAGCTGAGGAGAC 7348
 Db 4358 AGAGATTATGACATTTTGGCAGGACCGTGGACAACTGACATAGAAATCTGGA---AA 4414
 QY 7349 CTTAGTAGAGCAAGCTCATGCTCTTTAATCTCTATATGTTTCAAAACCAACACCCCATC 7408
 Db 4415 AGAATTGGGAAGACCCAGCTCTTTTAACTTGTACATGTTTAGGAATAAACAACCCGTCA 4474
 QY 7409 AGCAATCCAGGACATCAAGTCTCATGACAGATGTTAAAGCTGGCTGCTCACTTT 7468
 Db 4475 AGCAAAATCCATGACGTTGAAGATTATCATCAGATGTTAAACAGCTGGCTGGTGCATTT 4534
 QY 7469 GGATTCAGCTACACAGCTGATCCCTGTTTATCCCAAAACAGACATGATGATGCAATGAA 7528
 Db 4535 GGTTCATCTGACAAATGCTATCTGATGATCCCTGTTCCCAAAATTTGATTTACAGAA 4594
 QY 7529 CCTCTCTACAGCTCTCCACACAGATGAAACCGAGTGGGACAAACAGCAAGTCT 7588

Db 4595 CTTCTTTACGA-----ACTTGTGAAGAGTTCAGCAGTGGGATGATATACCGCCC 4642
 QY 7589 ATCTCTGGGTACAGTGTGAAGTACAGAGCACTCAAGGCTTTGTCACTTTAGAACGG 7648
 Db 4643 ATCTCTGGAGTCCAGCAGCAAGTGGCGCGCAGGCCAAGGCTTCTCTGCTGGGAAG 4702
 QY 7649 TTTTACCAGCTCTATGCTCTCCAAATCACCAGCTGCGCAGCAGGCTCCAAAGACCAAGAAG 7708
 Db 4703 ATGCCCGAGGTGACAGTGGCGCGCGCGCGCGC---GGCGGCAAGTCTCTGCTGTGG 4759
 QY 7709 TTTTCATCCAGCGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGGATGSCCGA 7768
 Db 4760 TTTCCCAAGCTCAAGTCTGCTGATCGGCAAGGGCGCTCATGCTGCGCGCTCAGCCAGGCGCG 4819
 QY 7769 GTGACCACAGACATCACTCAGTGTGGCAATGAGATGGGCGAAGGTTGCTGCATCTTG 7828
 Db 4820 GTGCAGACCAACGTCTCAACATCGCCACAGAGACTGCATCAAGTGGCGCGCTGCTC 4879
 QY 7829 AACCATGCCACTACCTAGAGAACTTCACATTCACCATTCATGGGTGGATACCCATTAC 7888
 Db 4880 AACAAAGCTTCTACCTGGAGAACTTCACATTCAGGCGCAAGGACAGCAGCTAC 4939
 QY 7889 TTTGTGAACCAAGACCTTTCAGAGGTGACCTGSCCATCTCGGCTCAGTGGGGCGG 7948
 Db 4940 TTTATCAAGACCAACCGCCGAGAGCACTGGGACGCTGGGTTGACAGCGCGCG 4999
 QY 7949 CGAACCTCTGGAGATGGGGTCAAGCTCACTGTGTCCAGATCAACACAGTACTTAATGGC 8008
 Db 5000 AAGCGCTGGAGAACGCACTCACTGACGCTGTCGAGTCCACCGTGGTGGTGAACGGC 5059
 QY 8009 AGGACTAGAGCTTACAGACATCCAGCTCCAGT 8042
 Db 5060 AGGACGCGAGGTTTCGCGGACGTTGAGATGACGT 5093

RESULT 4
 AY413477 3190 bp DNA linear GSS 12-DEC-2003
 LOCUS Mus musculus HCM4903 gene, VIRTUAL TRANSCRIPT, partial sequence,
 DEFINITION genomic survey sequence.
 AY413477
 VERSION AY413477.1 GI:39769439
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 REFERENCE 1 (bases 1 to 3190)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,S., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 Science 302 (5652), 1960-1963 (2003)
 JOURNAL 14671302
 PUBMED 2 (bases 1 to 3190)
 REFERENCE Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,S., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1. 3190
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 <1..>3190
 /locus_tag="HCM4903"
 gene

ORIGIN

Query Match 28.1%; Score 2347.8; DB 29; Length 3190;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 2612; Conservative 0; Mismatches 474; Indels 48; Gaps 2;

Qy	2161	CACCGGCTTGCAGCTGTGACCAAGCTGGAGTGGACAGACTGTCTATCGAGATCTG	2220
Db	105	CATGGGACCTGCTGCAACCCCGCTACAAGGGGAGAGCTGCGAGGAAGAGATCTG	164
Qy	2221	TGCTGCCGACTGTGGTGGCCATGCGGTGCGTAGGGGCACTGCGCTGCGAGATGG	2280
Db	165	TGCCGCTGACTGTGGTGGCCATGCGGTGCGTAGGGGCACTGCGCTGCGAGATGG	224
Qy	2281	CTGATGGGGGCACTGCGACCAAGCGGGCTGCAACCCGCTGTGCCGAGCATGGAC	2340
Db	225	TTGATGGGGGCGCATGCGACCAAGCGGGCTGCAACCCGCTGTGCCAACAACGGGAC	284
Qy	2341	CTGCGGCGAGGCAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG	2400
Db	285	CTGCGGCGAGGCAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG	342
Qy	2401	TCATCTCTGGATAGGCTAGTAAAGAGGGTGGCCCTGGGTCTGTCAATGGCAACGGCAG	2460
Db	343	-----GNN	377
Qy	2461	ATGTACCTTAGACCTGAATGGTGGCACTGCGCTGCGAGCTGGGCTGGAGAGAGCTGG	2520
Db	378	NN	437
Qy	2521	CTGTGACACTTCCATGGAGACTGCTGCGGTGACAGCAAGACAAATGATGGAGATGGCT	2580
Db	438	NN	497
Qy	2581	GGTGGACTGATGGAACCTGACTGCTGCCAGCCCTGTGCCATATCAACCCGCTGTG	2640
Db	498	GGTGGACTGATGGAACCTGACTGCTGCCAGCCCTGTGCCATATCAACCCGCTGTG	557
Qy	2641	CCTTGGCTCCCTAAACCTGTGACATCATCCAGAGACACAGGTCCCTGTGTACAGCA	2700
Db	558	CTAGGCTCCCTGACCCCTGACATCATCCAGAGACACAGGCCCTGTATCCAGCA	617
Qy	2701	GAACCTACACTCTTATGACCCCATCAAGTTCCTCGTGGGAGGAGCAGACGACAT	2760
Db	618	GAACCTGAACCTCTACGATCGAATCAAGTTCCTCGTGGGAGGAGCAGACACAG	677
Qy	2761	AATCCCGGGGAGAACCTTTGATGGAGGGCATGCTTGTGTTATTCGTGGGCAAGTAT	2820
Db	678	CATCCCGGGGAGAACCTTTGATGGAGGGCATGCTTGTGTTATTCGTGGGCAAGTAT	737
Qy	2821	GACATCAGATGGAAACCCCTGCTGTGTGTAACATCAGTTTGTCAATAACCTCTCT	2880
Db	738	GACGTGAGATGGAAACCCCTGCTGTGTGTAACATCAGTTTGTCAATAACCTCTCT	797
Qy	2881	TGGATATACATCAGAGGCAAGATGGCAGCTTGAATGGTGAACAAATGGGGCATCTC	2940
Db	798	TGGATATACATCAGAGGCAAGATGGCAGCTTGAATGGTGAACAAATGGGGCATCTC	857
Qy	2941	CATCATCCTGCGGTGGAGGGGACCTTTTCATCACAGGAGCACACCTGTGGCTGCC	3000
Db	858	CATCATCCTGAGATTGGAAGGGGACCTTTTCATCACAGGAGCACACCTGTGGCTACC	917
Qy	3001	ATGGGATCGCTTCTTGTGATGAAACCATCATCATGAGACATGAGGAGAATGAGATCC	3060
Db	918	CTGGGATCGCTTCTTGTGATGAAACCATCATCATGAGACATGAGGAGAATGAGATCC	977
Qy	3061	CAGCTGTGACTGAGCAATTTTCCCGCGCCCAACCCAGTGTCTCTCCATCCCACTGAC	3120
Db	978	CAGCTGTGACTGAGCAATTTTCCCGCGCCCAACCCAGTGTCTCTCCATCCCACTGAC	1037
Qy	3121	GTCTTTCGCCAGTCTGTGTGAGAGAAAGCCCAATTTGCCAGGATAGCTGGAGTGCC	4260
Db	1038	ATCCTTTCGCCAGTCTGTGTGAGAGAAAGCCCAATTTGCCAGGATAGCTGGAGTGCC	1097

Qy	3181	GGAGGAATCTCTATCTCTGGCTCAAGATGAGGTGAGCTTACCTGAGCAGCGGACCCC	3240
Db	1098	AGAGGAATCTCTATCTCTGGCTCAAGATGAGGTGAGCTTACCTGAGCAGCGGACCCC	1157
Qy	3241	TGGCTACAAATCTGTCTTCTGAGATCAGCTTCAACCCCGGACCATCCCTTCAACTCAT	3300
Db	1158	TGGCTATAGTCTGTCTTCTGAGATCAGTCTCACACCCCGGACCATCCCTTCAACTCAT	1217
Qy	3301	GAGGTGACCTCATGTTAGCGGTGGAGGGCGGCTCTTCAAGAAAGTGGTTCGCTGAGC	3360
Db	1218	GAGGTGACCTCATGTTAGCGGTGGAGGGCGGCTCTTCAAGAAAGTGGTTCGCGGAGC	1277
Qy	3361	CCGAGACCTGTCTCTATTTTCAATTTGGGAACAAGACAGACGCTCTACAAAGAGTGT	3420
Db	1278	TCCTGACCTGTCTCTATTTTCAATTTGGGAACAAGACAGACGCTCTACAAAGAGTGT	1337
Qy	3421	TGGGCTTTTCAAGAGCTTTTTCGTTGGTTATGAATATGAATCTTCCGAGATCTAAT	3480
Db	1338	TGGACTCTCAGAAGCTTTTTCGTTGGTTATGAATATGAATCTTCCGAGATCTAAT	1397
Qy	3481	CTGTGGGAAAAAAGAACCAAGTCTGAGGGCTATGAAATTTGACGCTTCAAGCTGG	3540
Db	1398	CTGTGGGAAAAAAGAACCAAGTCTGAGGGCTATGAAATTTGACGCTTCAAGCTGG	1457
Qy	3541	AGGATGGAGCTTAGACAAACATCATGCTTCAACATTTCAAGTGGTATCTTGCAAAAG	3600
Db	1458	GGGCTGGAGTCTGGATAAGACCAATGCTTCAACATTTCAAGTGGTATCTTGCAAAAG	1517
Qy	3601	GATGGGGAGAACCAATTTTGTCTCAGACGCTTCTGTCTATGGGAGCATATGGGAA	3660
Db	1518	GATGGGGAGAACCAATTTTGTCTCAGACGCTTCTGTCTATGGGAGCATATGGGAA	1577
Qy	3661	TGGGCGCGGAGAGACATCTCTGCTCCCGAGCTGCAACCGGCTTGTGAGGCAACAGCT	3720
Db	1578	TGGGCGCGGAGAGAGTATCTCTGCTCCCGAGCTGCAACCGGCTTGTGAGGCAACAGCT	1637
Qy	3721	CCTGGCCCCAGTGGCCCTCAGCTGTGGCTCTGAGGGAGCCCTCTATGTGGGTGATTTCA	3780
Db	1638	CCTGGCCCCAGTGGCCCTCAGCTGTGGCTCTGAGGGAGCCCTCTATGTGGGTGATTTCA	1697
Qy	3781	CTACATTTAGAGAGTCTTCCCTCTGGAATGTGCAACCATCTTAGAGCTGAGGAATA	3840
Db	1698	CTACATTTAGAGAGTCTTCCCTCTGGAATGTGCAACCATCTTAGAGATG-----	1749
Qy	3841	AGATTTTCAAGATCAGTCCAGCACACAAATATCTCTGCGGAGCAGACCCCATGAG	3900
Db	1750	-----AGTCAAGCCAGCACACAAATATCTCTGCGGAGCAGACCCCATGAG	1796
Qy	3901	TGGGCGCGTCTTCTTCTGAGCAGACAGCGGGGCTTTTAAATCAAGTCACTGT	3960
Db	1797	TGGGCGCGTCTTCTTCTGAGCAGACAGCGGGGCTTTTAAATCAAGTCACTGT	1856
Qy	3961	GGTGTGAAGGACCTTTCTCAAGAACTCTGAGGTGGTTGCGGGGACAGTGACCAAGTGCCT	4020
Db	1857	AGTGTGAAGGACCTTTCTGAGAACTCTGAGGTGGTAGCAGGAGTGGTACCAAGTGCCT	1916
Qy	4021	CCCCCTTTGATGACACTGCTGCGGGATGTTGGGAAGGCGACAGAACCACTCACAA	4080
Db	1917	CCCCCTTTGATGATACCCGCTGCGGAGTGGTGGGAAGGCGACAGAACCACTCACAA	1976
Qy	4081	TCCAGGGGTATTACAGTGGCAAGTTTGGGCTCATCTCTCTGATGGAGCAGCATGAT	4140
Db	1977	CCCCAGGGGATTACAGTGGCAAGTTTGGGCTCATCTTATTTCTGAGCGGACCATGAT	2036
Qy	4141	CAGACGCAATCAGATGGGATCATCTCACCCCTGCTGGCTCTAATGATCTCACATC	4200
Db	2037	CAGACGTTGATCAAAATGGAATCATCTCCACTTTGCTGGGCTTCAATGATCTCACCTC	2096
Qy	4201	AGCCCGGCGCTCAGCTGTGATCTCTGTCATGATTTTCCAGGTAGATGGAGTGCC	4260
Db	2097	GECCAGGCGCTCAGCTGTGACTCCGTATGGAGATTTCTCAGGTTCGCTGGAGTGCC	2156

QY 4261 CACAGACTTAGCCATCAACCCATGACAACTTATGTCTCTCGACCAAAATGTGGT 4320
 Db 2157 CACAGACTTAGCCATCAACCCATGATTAATCTCTATGTCTCTCGACCAAAATGTGGT 2216
 QY 4321 CTTGCAAACTCTGAAACACACAGGTGGCATTGCGCGGAGGCCCATGCACTGGCA 4380
 Db 2217 CTTGCAAACTCTGAAACACACAGGTGCGATTGCTGCGGAGGCCCATGCACTGGCA 2276
 QY 4381 GGTCCCTGGCAATGACATCTTCTGTAAGCAAGTGGCCATCCAGCAACCTGGAGTC 4440
 Db 2277 GGTCCCTGGCAATGACATCTTCTGTAAGCAAGTGGCCATCCAGCAACCTGGAGTC 2336
 QY 4441 AGCCACCGCTTGGCTGTTTACACAAATGGGTCCTGATATGCTGAGACTGATGAGAA 4500
 Db 2337 AGCCACCGCTTGGCTGTTTACACAAATGGGTCCTGATATGCTGAGACTGATGAGAA 2396
 QY 4501 AAAGATCAACCGCATCAGGAGTCAACCACTAGTGGAGAGATCTCACTGTTGCTGGGC 4560
 Db 2397 AAAGATCAACCGCATCAGGAGTCAACCACTAGTGGAGAGATCTCACTGTTGCTGGGC 2456
 QY 4561 CCCAGTGGCTGACTGTAATAATGATGCCAATGATGTTTCTCGAGAGCATGG 4620
 Db 2457 CCCAGTGGCTGACTGTAATAATGATGCCAATGATGTTTCTCGAGAGCATGG 2516
 QY 4621 TTATGCCAAGGATGCAAGTAAATACCCCATCTTCTTGGCTGTGTGCTGATGGGA 4680
 Db 2517 TTATGCCAAGGATGCAAGTAAATACCCCATCTTCTTGGCTGTGTGCTGATGGGA 2576
 QY 4681 GCTTACGTGGCACTCTGGGAACATCGAATTCGGTTATCCGGAAGCAAGCCTT 4740
 Db 2577 GCTTACGTGGCACTCTGGGAACATCGAATTCGGTTATCCGGAAGCAAGCCTT 2636
 QY 4741 CCTCAACCCAGCAATGATGATGCTCTTCAACCAATGACAGAGCTCTATCTGT 4800
 Db 2637 CTTGAACCTTCAGAACATGATGAGCTATCTCTCCCATCGACAGAGCTGTACTCTT 2696
 QY 4801 TGATACCCCGCAAGCACTGTACACCCAAAGCTGCGCCACAGGAGACTCTGTGACAA 4860
 Db 2697 TGATACCCCGCAAGCACTGTACACCCAAAGCTGCGCCACAGGAGACTCTGTGACAA 2756
 QY 4861 CTTCACTACCTGGGACCGGCACATCAGTCACTCAGAGTGTGACACACAGGACAGCA 5040
 Db 2757 CTTCACTACAGGAGCGGACATCAGATATCAGGACAAATGGAACATGAT 2816
 QY 4921 AAATGTCCGCGAGACTCTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGT 4980
 Db 2817 GAAGTCCGCGAGACTCTACCGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGT 2876
 QY 4981 GTACTGGGTGACCTGGGACCGACAGTCACTCAGAGTGTGACACACAGGACAGCA 5040
 Db 2877 ATACTGGGTGACCTGGGACCGACAGTCACTCAGAGTGTGACACACAGGACAGCA 2936
 QY 5041 GTTGCCATGATGATACATACCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAA 5100
 Db 2937 GCTAGCCATGATGATACATACCAATTCGGGCTTCTGGCAACCAAAAGCAATGAAA 2996
 QY 5101 CGGATGACAACTTTATGAGTACGACAGCTTGGCGGCTGACAAATGTGACCTTCC 5160
 Db 2997 CGGATGACAACTTTATGAGTACGACAGCTTGGCGGCTGACAAATGTGACCTTCC 3056
 QY 5161 TACTGGCCAGTGGAGCTTTCGAGTGTATACAGACAGTTCAGTGCATGCTCCAGTGA 5220
 Db 3057 AACTGGCCAGTGGAGCTTTCGAGTGTATACAGACAGTTCAGTGCATGCTCCAGTGA 3116
 QY 5221 GACTCCAGCAAGATGATGACCAATACCAACCACTGTGCTGCTCAGGCGCTTCTTA 5280
 Db 3117 GACTCCAGCAAGATGATGACCAATACCAACCACTGTGCTGCTCAGGCGCTTCTTA 3176
 QY 5281 CACTGCTGCAAG 5294
 Db 3177 CACCTGTTACAAG 3190

RESULT 5
 LOCUS AY405421 5069 bp DNA linear GSS 12-DEC-2003
 DEFINITION Pan troglodytes HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
 ACCESSION AY405421
 VERSION AY405421.1 GI:397611395
 KEYWORDS GSS
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 REFERENCE 1 (bases 1 to 5069)
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
 TITILE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
 FEATURES
 source 1. 5069
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 gene <1..5069
 /locus_tag="HCM2218"
 ORIGIN
 Query Match 27.3%; Score 2282.2; DB 29; Length 5069;
 Best Local Similarity 65.4%; Pred. No. 0;
 Matches 3340; Conservative 0; Mismatches 1720; Indels 45; Gaps 6;
 QY 2912 TTTGACTTGTGACAAATGCGGCACTCTCCATCATCTCGGTTGCGAGCGGCACCTTTC 2971
 Db 2 TTTGACTTGTGCGAAATGCTGGGCTCTCTAACTTTGTAATTGAAACGATCCCAATC 61
 QY 2972 ATCACAGGAGCACACCTGTGGCTGCCATGGATCGCTTTCTTGTATGGAACCATC 3031
 Db 62 CTCACCGATATCATCTGTGTGGATTCATGGAATGCTTTTATGTGATGATACCTTA 121
 QY 3032 ATCATGACATGAGGAGAAATGAGATTCACGCTGTGACCTGAGCAATTTTCCCGCCCC 3091
 Db 122 GTCATGAAGAAGAAGAAGATGACATTCACGAGCTGTGATCTGAGTGGATTCGTGAGGCCA 181
 QY 3092 AACCCAGTCGTCTCTCCATCCCACTGAGCTCTTCCGCCAGTCTCTGTGAGAGAAAGGC 3151
 Db 182 AATCCCATCATGTGTGCATCACTTTTATCCACCTTTTTCAGATCTTCTCTGAGACAGT 241
 QY 3152 CCCATTGTGCGGAAATTCAGGCTTTTCAGGAGGAAATCTCTATCTCTGCTGCAAGATG 3211
 Db 242 CCCATCAATTCGGAACACNNNNCTCCACGAGGAAACTACAAATCCAGGAACAGATTG 301
 QY 3212 AGGCTGAGTACCTGAGCAGCGGCCCTGGCTTACAAATCTGTCTCTGAGATCAGCTC 3271
 Db 302 AAATCTCTCTACTTGTGAGTTCAGAGCTGCGAGGTATAGTCAGTCTCTCAAGATCACCATG 361
 QY 3272 ACCACCCGACCATCCCTTCAACCTCATGAAGTGCACCTCATGTTAGCGGTGAGGGC 3331
 Db 362 ACCAGTCTATTATTCCTATTATTAATGAAGTTTCATCTGATGTTAGTCTAGTAGGA 421
 QY 3332 CGCTCTTCAGGAAGTGGTTCGCTGCGAGCCCAACCTGTCCTATTATTCATTGGGAC 3391

Db	422	AGACTCTTCCAAAGTGGTTTCCTGCTCACCACCACTGGCCCTATATCTTTCATATGGGAT	481
Qy	3392	AAGACAGACTCTACACACAGAGGTTTGGGCTTTCAGAACCTTTGTTCCGTTGGGT	3451
Db	482	AAACAGATGATATATATCAAGAGTCTATGGTCTGTCTGAAGCTGTGTGTGAGTTGGA	541
Qy	3452	TATGAATATGAATCTCTGCCAGATCTAATCCTTGGGAAAAAAGAAACACAGTGTGTCAG	3511
Db	542	TATGAGTATGAGTCGTGTTTGGACCTGACTCTCTGTGGAAAAAGAGGACTGCCATCTCTG	601
Qy	3512	GGCTATGAATATGACGGCTCAAGCTTGGAGATGGAGCTAGACAAACATATGCCCCC	3571
Db	602	GGCTATGAATGAGTGGCTCCACATGGGTGGCTGGACATATAGATAAACATCAGTGTG	661
Qy	3572	AACATTCAAAAGTGGTATCTCTGCAAAAGGAAATGGGGAGAACACAGTTTGTCTCAGCAG	3631
Db	662	GATGTACAGAACGGTATACCTGTACAGGAAACGGGGAAACACAGTTTCTCTCCAGCAG	721
Qy	3632	CTCTCTGTATTTGGAGATCATGTGGCAATGGGCGCGGGAGAGATCTCTCTGCCCCAGC	3691
Db	722	CTCTCCAGTCTGTAGTAGCATCATGGGCAATGGGCGGAGGCGCAGCATTTCTCTGCCCCAGT	781
Qy	3692	TGCAACGGCTTGTGACGGCAACAAGCTCTCTGCCCCCAGTGGCCCTCACCCTGTGGCTCT	3751
Db	782	TGCAATGGTCAAGCTGATGTATCAAGTTACTTGGCCCCAGTGGCGGTAGCTTGTGGATC	841
Qy	3752	GACGGAGCTCTATGTGGTGTATTTCACTACATTAGAAGATCTTCCCTCTGGAAT	3811
Db	842	GACGGAGCTGTGACGTAGGCGATTTCACTATGTGGCGGATATTTCCCTCTGGAAT	901
Qy	3812	GTCAACCAATCTTAGAGCTGAGGATTAAGATTTTCAACATAGTCACAGTCCAGCACAC	3871
Db	902	GTAAACAGTGTCTTAGAACTAAGCA-----GCAACCCAGCTCAT	940
Qy	3872	AAATACTACCTGGCCACAGACCCCATGAGTGGGCGGCTCTTCTCTTCTGACACCAACAGC	3931
Db	941	AGATACTACCTTGCAACGGATCCAGTCACAGGAGATTTGTACGTTTCTGACACAAACACC	1000
Qy	3932	CGCGGGCTTTTAAATCAAGTCCACTGTGGTGGTGAAGACCTTGTCAAGAACTCTGAG	3991
Db	1001	CGCAGAAATTTATCGCCCCAAGTCACTTACGGGGGCAAAAGACTTGAATAAAATGCAGAA	1060
Qy	3992	GTGGTTCGGGGACAGTGTACAGTGCCTCCCTTTTGTATGACACTCGCTCGGGGATGT	4051
Db	1061	GTGTGCGGGGACAGGGAGAGTGCCTTCCGTTTGAAGGCGAGATGTGGGGATGA	1120
Qy	4052	GGGAAGCCACAGAAGCCACACTCAACAAATCCAGGGGTATTACAGTGGACAAAGTTTGG	4111
Db	1121	GGGAAGGCCCTGGAAGCCACACTCATGATGTCCTCAAGGAATGGCAGTTGATGAAGATGA	1180
Qy	4112	CTGATCTACTTCTGTGATGGACCATGATCAGACGATCGATCAGATGGATCATCTCC	4171
Db	1181	TTAATCTACTTGTGTATGGAAACCATGATTAGAAAGTTGACCAAAATGGAATCATATCA	1240
Qy	4172	ACCTGCTCGGCTCTAATGATCTCAATCAGCCGGCCACTCAGCTGTGATTTCTGTCATG	4231
Db	1241	ACTCTTCTGGCTCTAACGATTTGACTTTCAGCCAGACCTTTAACTTGTGACACCAGCATG	1300
Qy	4232	GATATTTCCAGATGATGAGTGGAGTGGCCACAGACTTAGCCATCAACCCATGAGCAAC	4291
Db	1301	CACATCAGCCAGGTACTCTGGAATGGCCCACTGACCTAGCCATTAAACCTTAGGATAAC	1360
Qy	4292	TCACTTTATGCTCTCGACAAACATGTGGTCTCTGCAAAATCTCTGAAACACCAAGGTGGC	4351
Db	1361	TCCATTATGCTCTGATANTATACGTAGTTTACAGATCACTGAAATGCTCAAGTTCCG	1420
Qy	4352	ATTGTGCGGGAGGCCATGCACTGCCAGTCCCTGGCATTTGACCACTTCTCTGTGTAAGC	4411
Db	1421	ATTGTGCTGGAGCGGCCATGCACTGTCAAGTTTCCCGAGTGG---AATATCTGTGGG	1477
Qy	4412	ARGGTGGCCATCCACGCAACCTGGAGTCAAGCCGCTTGGCTGTTTTCACACATGGG	4471

Db	1478	AAGCAGCAGTGCAGACAAACACTGGAATCAGCCACTGCCTATGCTGTCTTACAGTGGG	1537
Qy	4472	GTCTGTATATTTGCTGACACTGATGAGAAAAGATCAACCCATCAGCAGGTGACCACT	4531
Db	1538	GTCTGTATATTTACTGAAAACCTGATGAGAAAATTAACCGGATAGCAGGTCAACA	1597
Qy	4532	AGTGAGAGATCTCACTCGTTGCTGGGCCCCCAGTGGCTGTGACTGTAAATAATGATGCC	4591
Db	1598	GATGAGAAAATCTCTTGTAGTGGCCGGAATACCTTCAGAGTGTGACTGCAAAAATGATGCC	1657
Qy	4592	AACTGTGATGTTTCTTGGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATACCCCA	4651
Db	1658	AACTGTGACTGTTACAGAGTGGAGATGGCTACGCCAAGGATGCCAAACTCAGTGGCCCCA	1717
Qy	4652	TCCTCTTGGGTGTGTGCTGATGGGAGCTCTACGTGGCGGACCTTGGGAAACATCCGA	4711
Db	1718	TCCTCTTGGGTGTCTTCCAGATGGTACACTGTATATTTGAGATCTAGGGAATATCCGG	1777
Qy	4712	ATTGGTTTATCCGGAAGAACAGCCCTTCTCAACCCCAAGACATGATGTAGCTGTCT	4771
Db	1778	ATCCGGGTGTGTCAAGAATAAGCCCTTTACTTAATCTATGAACTTCTATGAAATTGCG	1837
Qy	4772	TCACCAATTGACAGGAGCTCTATCTGTTTGATACCCAGCGCAAGCCTCTGACACCCAA	4831
Db	1838	TCTCCACTGATCAAGAACTCTACATCTTTCAGATCATGTGTACTCACCATATACTGTA	1897
Qy	4832	AGCTGCCCAAGAGAGACTACTCTGACAACTTCACTTACCTGAGGAGCGGACATCAACA	4891
Db	1898	AGTTTAGTCACTGGTGATTACCTTTTACAATTTTAGCTACAGCAATGACAATGATTTACT	1957
Qy	4892	CTCATCAGACAAACATGGCAACATGTAATGTCCGCGGAGACTCTACTGGGATGCC	4951
Db	1958	GCTGTGACAGACAGCAATGGCAACACCTTTAGAAATTAGCGGGACCCCAATCGCATGCCA	2017
Qy	4952	CTCTGGCTGGTGGTCCAGATGGCAGGTGACTGGGTGACCATGGGCAACCAAGTGCAC	5011
Db	2018	GTTCAGTGGTGTCTCCCGATAACCAAGTGATATGGTTTGACAAATAGGAACAAATGCGATG	2077
Qy	5012	CTCAAGATGTGACCAACAAGGACACGAGTTGGCCATGATGACATCCATGGCAATTC	5071
Db	2078	TTGAAAGCATGACTGTCTCAAGGACTGGAATTTAGTTTNNGTTTACTTANNNNNNNNNNN	2137
Qy	5072	GGCTTCTGGCAACCAAAAGCAATGMAAOCGATGGACAAACATTTATGATGACACAGC	5131
Db	2138	NN	2197
Qy	5132	TTTGCCGCCCTGACAAATGTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGA	5191
Db	2198	GAAAGTCTCTGACAAATGTTTACGTTTCCAACTGGAGTGGTCAACAAACCTGCATGGGAC	2257
Qy	5192	ACAGACACTTTCAGTGCATGTCAGGTAGAGACTCCAGC---AAGGATGATGTACACATA	5248
Db	2258	ATGGCAAGGCTATCAGAGTGACATTTAGTGCATCTAGCCGAGAGGAAGATGTGAGCATC	2317
Qy	5249	ACCCCAACCTGTCTGCTCAGGGCCCTTCTACACACTGTGTCGAAGACCAAGTCCGGAAC	5308
Db	2318	ACTTCAATCTGCTCTCGATCGANTTCTTCTACACCATGTTTCAAGATCAGTTAAGAAC	2377
Qy	5309	AGCTACTACATCGGGGCCGATGGCTCTTGGCGCTGCTGGCGCAACCGGATGAGAGTG	5368
Db	2378	AGCTACAGATTTGTTATGACGGCTCCCTCAGAAATTTACTACGCGAGCGGCTGAGCTCA	2437
Qy	5369	CGCTGACAGCTGAGCCCCACTTGTGGCTGGCACCGCTCAACCCACCGTGGGCAAGAGG	5428
Db	2438	CACCTACCAACAGAGCCGCAATTTCTGGTGGCTGCTGCTAATCCAAACGTTGCCAANAAG	2497
Qy	5429	AATGTACGCTGCCATGCAACAGGCTCAACTGTGTGGAGTGGCGGACCGCAAGAG	5488
Db	2498	AACATGAGCTTGCCTGTGTGAGAACGGTCAAAATTTGTGGAAATGAGATTCGGAAGAG	2557
Qy	5489	CAGGCTCGGGCCAGGTCACTGTCTTGGCGCGGCTGGCGTGCAACCCGAAATCTC	5548
Db	2558	CAAGCCCAAGGGAAGTCAATGTCTTTGGCCGCAAGCTCAGGGTTAATGGCAGAAACCTC	2617

QY 5549 CTATCTCGACTTTGATCGGTAAACGGCAGAGAAAGATCTATGATGACACCGCAG 5608
Db 2618 CTTTCAGTTGACTTTGATCGAACAACAGAAAGATCTATGACGACACCGTAA 2677
QY 5609 TTCACCCCTTCGGATTCTGTACGACAGGGGGCGGCCAGCCTCTGGTCAACCCAGCAGC 5668
Db 2678 TTTCTACTAGGATCGCTACGACAGCTCTGGGACCCGACTCTCTGGCTGCCAGCAGC 2737
QY 5669 AGGCTGAATGGTGTACAGTCACTATCCCTGGGGTTTACATGCTGGCATGCCAGAG 5728
Db 2738 AAGCTGATGGCCGTCAATGTCACTATTCATCCACAGGTCAAAATGCCAGCATCCAGCGA 2797
QY 5729 GGCATCATGTCTGAAGATGGATACGACCGAGCGGGCGGCATCACATCCAGGATCTTC 5788
Db 2798 GGCACCACTAGCGAAGTAGATTATGACGACAGGGAGGATCGTCTCGGGTCTTT 2857
QY 5789 GCTGATGGGAAGACATGGAGCTACACATCTTAGAAGTCCATGGTGTGCTACTACAC 5848
Db 2858 GCTGATGGTAAACATGGAGTTACACATATTTAGAAAAGTCCATGTTCTCTGCTTCAT 2917
QY 5849 AGCCAGAGGAGTATATCTTTGAGTTGCAAGAATGACGGCTCTCTTCTGTGAGGATG 5908
Db 2918 AGCCAGGGCAGTACATCTTCGAATACGATATGCGGACCGGCTGTCTGCCATCACCATG 2977
QY 5909 CCCAAGCTGGCGGAGACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATC 5968
Db 2978 CCAGTGTGGCTGCCACACCATGCGACCATCCGATCCATTTGGCTACTACCGACATA 3037
QY 5969 TATCAGCCCCCTGAGGGCAATGCTCAGTCAATACAGGACTTCTACTGAGGATGGGCACCTC 6028
Db 3038 TACAACCCCCGGAAGCAACGCTCCATCATCAGGACTACAAACGAGGAGGGGTGCTT 3097
QY 6029 CTTTACACCTTCTACCTGGGCACTGGCGGAGGTGATATACAAGTATGCGAAATGTCA 6088
Db 3098 CTGCAAAACAGCTTTCTTGGGTACAAGTCGGAGGGTCTTATTTCAAATACAGAGGAGACT 3157
QY 6089 AAGCTGCGAGAGCGCTCTATGACACCAACCAAGSTCAGTTTCACTTATGACGAGCGCA 6148
Db 3158 AGGCTCTCAGAAATTTATATGATAGCAGGAGTCACTTTTACCTTACGATGAACAGCA 3217
QY 6149 GGCATGCTGAGACCATCACTACAGATGAGGGTTCACCTGACCACTCGCTACCGT 5208
Db 3218 GAGTCTCTAAAGACAGTAAACCTCCAGAGTGACGGTTTATTTGCACCATTAGATACAG 3277
QY 6209 CAGATTGGGCCCCGTGATGACGACAGATCTCCGCTTCACTGAGGAAGGCATGTCAAC 6268
Db 3278 CAAATTTGCTCCCTGATTGACAGGAGATTTTCCGCTTTAGTGAAGATGGGTAAAT 3337
QY 6269 GCCGTTTGTACTACAACTATGACACAGCTTCCGGGTGACAGCATGAGGCTGTGATC 6328
Db 3338 GCAAGATTTGACTATAGCTATGACACAGCTTTCGAGTGACCAAGCATGAGGGTGTGATC 3397
QY 6329 AACGAGACCCACTGCCCCATTGATCTCTATCGCTATGATGATGTGTCAGGCAAGACAG 6388
Db 3398 AATGAAACTCCATGCTCATTTGATCTGATCAGTTTGTATGACATTTCTGGCAAGTTGAG 3457
QY 6389 CAGTTTGGGAAGTTGGTGTCACTTTACTATGACATTAACAGATCATACCAAGCTGTG 6448
Db 3458 CAGTTTGGAAAGTTGGAGTTATATATGATATTAACAGATCATTTTCTACAGCTGTN 3517
QY 6449 ATGACCCACACCAAGCATTTTGTATGATGCGAGGTGAAGAGTGCAGTATGATGATC 6508
Db 3518 ATGACCTATACGAGCACTTTGATGCTCATGGCCGTATCAAGGAGATTCAATATGATA 3577
QY 6509 TTCGCTCGCTCATGTACTGATGACCGTTCAGTATGATAAATGAGGGCGAGTAGTGAAG 6568
Db 3578 TTCAGGTCGCTCATGTACTGATTAACAATTCAGTATGATAAATGAGGTCGGGTAAACCAAG 3637
QY 6569 AAGGAGTGAAGTAGACCTAGCGCAATACCACTCGCTACTCTCTATGATGATGATGCT 6628
Db 3638 AGAGAGATTAAATAGGGCCCTTTGGCAACACCAACCAATATGCTTATGAATATGATGT 3697

QY 6629 GACGGCAGCTGACAGCAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTACGAC 6688
Db 3698 GATGACAGCTCCAAACAGTTTACTCAATGAAGAAGTAATGTGGCGGTACAACTACGAT 3757
QY 6689 CTCATATGGGAACCTTGCACTTACTGAGCCCTGGAAACAGTGCACGGCTTCACACCACTACGG 6748
Db 3758 CTGAATGGAAACCTTCCATTTACTTGAACCCCAAGTAAACAGTGGCGGTCTGACACCCCTTCG 3817
QY 6749 TATGACATCCGACCCGATCCTCGCTGGGTGACGCTGCAATACAAGATGATGAGAT 6808
Db 3818 TATGACCTGGAGACAGAACTCCTGACTGGGTGATGTTCAATATCGTTGGATGAAGAT 3877
QY 6809 GGCCTTCTGAGGACGCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCCCTGCTCATC 6868
Db 3878 GGTTCCTACGTCAAAGGGCACCGAAATCTTTGAATATAGCTCCAAGGGCTTCTAACT 3937
QY 6869 AAGCCCTAACACCGGGCTGGAGCTGGAGTGTGAGTACCGCTACGATGGCCTGGGGCGG 6928
Db 3938 CGAGTTTACAGTAAAGGACGTGGCTGGACAGTGANNNACCGTTATGACGGCCTGGGAAG 3997
QY 6929 CGCGTGTCCAGCAGACAGCCACAGCCACCTGCGAGTCTTCTATGACAGACCTGACC 6988
Db 3998 CGTGTNNAGCAAAACAGTCTAGGACAGCACTGCGAGTTTTTTATGCTGACTAACT 4057
QY 6989 AACCCCAACAGGTACCCACCTGTACAACACTCCAGCTCTGAGATCACTCCCTCTAC 7048
Db 4058 TATCCCACTAGGATTACTCATGCTACAACCAATCGAGTTTCAAGAAATTAACCTCCCTGTAN 4117
QY 7049 TACACTTGAAGACACCTCTTTGCCATGGAGCTGAGAGTGGTATGATGATTTTACATA 7108
Db 4118 NNTGATCTCCAAGACATCTTTTGGCCATGGAATCAGCAGTGGGATGAATTTCTATTT 4177
QY 7109 GCTTGTGACAAACATCGGAGACCCCTCTTCTGTCTTTAGTGGAAACAGGTTTGTATCAAG 7168
Db 4178 GCATCAGATAACACAGGACACCACTGGCTGTGTTTCAAGTAGCAATGGCTTATGCTGAAA 4237
QY 7169 CAAATCTGTACAGCCTATGGGGAGTCTACATGGATACCAACCCCACTTTCAGATC 7228
Db 4238 CAGATTGATACACTGCTATGTGAAATCTATTTGACTCTAATATGACTTTCACATG 4297
QY 7229 ATCATAGCTACCATGTGGCTCTATGATCCACTACCAAGCTTGTCCACATGGGCGCG 7288
Db 4298 GTAATGATTTTCAATGTGGCTGTATGACCCACTACCAATTAATCCATNNNNNNN 4357
QY 7289 CGAGATTATGATGTGTGGCGGACGCTGAGTAGCCCAAGCAGCAGCTGTGGAAGCAC 7348
Db 4358 NNN 4414
QY 7349 CTTTGTAGCAGCAACGTCATGCTTTTAAATCTCTATATGTTTCAAAAACAACACCCCATC 7408
Db 4415 AGAATCGGGAAGGACCCAGCTCTTTTAACTGTGATGTTTAGAATAACACCCCTGCA 7474
QY 7409 AGCAACTCCAGGACATCAAGTGTCTCATGACAGATGTTAACAGCTGGCTGCTACCTTT 7468
Db 4475 AGCAAAATCCATGACGTGAAAGATTACATACAGATGTTAACAGCTGGCTGGTGACATTT 4534
QY 7469 GGATTCAGCTACACACGTCATCCCTGTTATCCCAAAACAGACATGGATGCCATGAA 7528
Db 4535 GGTTTCATCTGCAATATGTTTCTCGATTCCTGATTCCTGTTTCCCAATTTGATTTAAACGAA 4594
QY 7529 CCCTCTCTACAGCTCATCCACACACAGATGAAACCGAGGATGGGACACCAAGTCT 7588
Db 4595 CTTTCTTACGA -----ACTTGTGAAGAGTCAGCAGTGGGATGATGATGCCGCC 4642
QY 7589 ATCTCTGGGGTACAGTGTGAAGTACAGAGCAGCTCAAGCCCTTTGTCACTCTAGAACGG 7648
Db 4643 ATCTCGAGTCCAGCAGAGTGGCGCGCAGGCCAAGCCCTTCTCTGCTGGGGAG 4702
QY 7649 TTTGACAGCTCTATGCTCCACAATCAACAGCTGCCAGAGGCTCCAAAGACCAAGAA 7708
Db 4703 ATGCCAGGTTGAGGTGAGCGCGCGCGCGCGCGC ---GGCGGCGAGTCTTGGCTGTG 4759
QY 7709 TTTGATCCAGCGGCTCAGTCTTTTGGCAAGGGGGTCAAGTTTGCCTTGAAGGATGGCCGA 7768

```

Db      4760 TTGCCACGGTCAAGTCGTCATCGGCAAGGGCGTCATGTCGCGCTCAGCCAGGCGCG 4819
Qy      7769 GTGACACAGACATCATCAGTGTGGCCAAATGAGATGGGGCAAGGGTTGTCCTCATCTTG 7828
Db      4820 GTGCAGACCAACGTCCTCAACATCGCCCAACGAGGACTGCATCAAGGTGGCGCGCTGCTC 4879
Qy      7829 AACCATGCCACTACCTAGAGAACTGCACCTTACCACTTATGATGGGTGGATACCCATTAC 7888
Db      4880 AACAAAGCCCTTACCTTGGAGACCTGCACCTTACCACTTACCACTTACCACTTACCACTTAC 4939
Qy      7889 TTTGTGAAACAGAGACCTTACAGAGGTGACCTGCGCCATCTCGGCGCTCAGTGGGGGCGG 7948
Db      4940 TTTATCAAGACCAACACCGCCGAGAGCGACCTGGCACGCTGGCGCTGACAGCGCGCGC 4999
Qy      7949 CGAACCTCGAGAAATGGGTCAAGCTCAGCTGCTGTCAGATGTCGCGATCAACAGTACTTATGGC 8008
Db      5000 AAGCGCTGGAGAAACGACATCAACGTCAGCTGTCGCGATGCCACCGTGGTGAACGGC 5059
Qy      8009 AGGAC 8013
Db      5060 AGGAC 5064

```

```

RESULT 6
AY405422      5087 bp DNA linear GSS 12-DEC-2003
LOCUS Mus musculus HCM2218 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY405422
VERSION AY405422.1 GI:39761396
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

```

```

REFERENCE
1 (bases 1 to 5087)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trics
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
2 (bases 1 to 5087)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..5087
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>5087
/locus_tag="HCM2218"

```

```

ORIGIN
Query Match 26.7%; Score 2232.8; DB 29; Length 5087;
Best Local Similarity 68.4%; Pred. No. 0;
Matches 3178; Conservative 0; Mismatches 1432; Indels 36; Gaps 5;
Qy 2912 TTTGACTTGGTGCAAAATGGCGGATCTCCATCATCTCGGGTTTCAGAGGGGACCTTTC 2971
Db 1 TTTGACTTGGTGCAAAATGGCGGATCTCTCTGACTTTGGTATTTGAGGTTCCCATTC 60
Qy 2972 ATCACAGAGGACACCCCTGTGGTGCCATGGATGGTCTCTTTGTGATGGAACCATC 3031

```

```

Db      61 CTCACCTCAGTACCACACTGTGTGGATTCCTTGAATGCTTTTATGTGATGGATACCTT 120
Qy      3032 ATCATGAGACATAGAGGAATGAGATTCCTAGCTGTGACCTGAGCAATTTTGGCCGCCCC 3091
Db      121 GTCATGAAGAAGAGAGAGAGACATTCCTCAGCTGTGACCTCAGTGGCTTTGTGAGGCCA 180
Qy      3092 AACCCAGTCTCTCTCCATGCCACTGACGCTCTTCCAGCTCTGCGAGCTCTGTGAGAGAAAGC 3151
Db      181 AGTCCCATCATTTGTCTTTCACCGTTATCCACCTTCTCAGGTCTTCCCTTGAGGACAGC 240
Qy      3152 CCCATTGTGCGGAAATTCAGGCTTTGAGGAGGAAATCTCTATCTCTGGCTGCAAGATG 3211
Db      241 CCCATCATCCCGAGACACAGCTGCTCATGAAGAACAACAATTCAGGAACAGATTG 300
Qy      3212 AGGCTGACCTACCTGAGCAGCGGACCCCTGGCTACAAATCTGTCTCTGAGGATCAGCCTC 3271
Db      301 AAATCTTCTTACCTGAGTTCAGAGCGGACAGGTACAAGTCAGTTCATTAAGATTACCATG 360
Qy      3272 ACCCACCGGACCATCCCTTCAACCTCATGAAGGTGCACTCATGTGACGCTGAGGCG 3331
Db      361 ACCCAGGCGGTACATACCGTTTAACTCATGAAGTCCATCTGATGTGCGCTGTTGG 420
Qy      3332 CGCCTCTTCAGGAAGTGTTCGTCAGCCCGGACCCAGACCTGTCTTATTTTCAATTGGGAC 3391
Db      421 AGACTCTTCCAGAAAGTGTTCCTGCTCGCCAAACCTTGGCTTACACGTTTCTCTGGAT 480
Qy      3392 AAGACAGACGCTACAAACAGAGGTGTTCGGCTTTCAGAAAGCTTTGTTTCCGTGGGT 3451
Db      481 AAGACGAGCGCATATAATCAGAAAGTCTACGCTTGTGAGGAGGATGTGTCCGTCGGA 540
Qy      3452 TATGAATATGAATCTGCCAGATTAATCTGTGGGAAAAAAGAACAAACAGTGTGCGAG 3511
Db      541 TACGAGTACGAGTCTGTGCTTGGACCTGACTCTCTGGGAAAAAGAGACTGCGCTTTGCAA 600
Qy      3512 GGCTATGAATGAGCGCTCGAAGCTTGGAGATGGAGCCTTAGACAAACATCATGCCCTC 3571
Db      601 GGCTATGATGGATGCTTCGAAACATGGCGGCTGGAGCTTGGGACAGACACCATGACTG 660
Qy      3572 AACATTAAAGTGGTATCTGTCACAAAGGAAATGGGGAGAACCAAGTTTGTCTCTCAGCAG 3631
Db      661 GAGCTTCAAGACGGTATATCTATACAAGAAATGGAGAAATCAGTTTCTCTCAGCAG 720
Qy      3632 CCTCTCTCATTTGGAGCATCATGGGCAATGGGGCCGAGAGAGCATCTCTGCCCCAGC 3691
Db      721 CCTCCGGTGGTCAGCAGCATCATGGTAAATGGTCGGAGCGTAGCATCTCATGCCAAGT 780
Qy      3692 TGCAACGCGCTTGTGACGGCAACAAGCTCTGGCCCGCAGTGGCCCTCACCTGTGGCTCT 3751
Db      781 TGCATGTCTAAGCTGACGGGAACAACCTCTGGCACCCGTCGCTGCTGTGGGATC 840
Qy      3752 GACGGGACCTCTATGTGGGTGATTTCAACTACATTAGAGAGATCTTCCCTCTGGAAT 3811
Db      841 GACGGCAGTATACCTAGGGGATTTCAATTTACGTCCGGCGGATATCCCGCTCTGGGAAT 900
Qy      3812 GTCACCAACATCTAGCTGAGGATTAAGATTTACACATAGTCAGAGTCCAGCACAC 3871
Db      901 GTGCAAGTGTTTTGA-----ACTAAGTAGCAACCCAGCTCAC 939
Qy      3872 AAATACTACCTGGCCACAGACCCCATGATGAGTGGGGCGCTTCTCTTCTTGACAGCAACAGC 3931
Db      940 AGATACTACCTGGCTACGGACCCAGTCACCGAGATTTGTACGTCTCTGTACTAACACC 999
Qy      3932 CGGGGGGTCTTTAAATCAAGTCCACTGTGTGGTGAAGGACCTTGTCAAGAACTCTGAG 3991
Db      1000 CGCAAAATCTATCGGCCGAAATCACTACGGGAGCCAAAGACCTGACTTAAAGACGCTGAA 1059
Qy      3992 GTGGTTGGGGGACAGAGTGCAGTGCCTCCCTTTTGTATGACACTCGCTCGGGGATGGT 4051
Db      1060 GTGGTGGAGGACCGGGGAACAGTGCCTTCCCTTTGACGAGGCCAGGTGTGGGATGGA 1119
Qy      4052 GGGAGGGCCACAGAAAGCCACACTCACCAATCCAGGGGTATTACAGTGGACAAGTTTGGG 4111

```

Db 1120 GGCAAGGCTGTGGAAGCAACGCTCATGTAGTCCCAAGGAATGCAATCGATAAGNACGGA 1179
Qy 4112 CTGATCTACTCTGTGGATGGCAACATGATCAGACGGATCGATCAGAAATGGGATCATCTCC 4171
Db 1180 CTGATCTACTCTGTGGATGGCAACATGATCAGAAAGGTTGATCAAAATGGAATCATATCA 1239
Qy 4172 ACCGTCTCGGCTCTAATGATCTACATCAGCCCGGCCACTCAGCTGTGATCTGTCAATG 4231
Db 1240 ACTCTCTGGCTCCACAGACCTCAGTCACTCGACCTTGACCTTTAACCTTGATAGCATG 1299
Qy 4232 GATATTTCCAGGTAAGATGGAGTGGCCCAAGATGACCAATGACCAATGGAACAAC 4291
Db 1300 CATATCAGCCAG-----CTGGAATGGCCCACTGACCTCGCATCAACCCCATGGATAAC 1353
Qy 4292 TCACCTTATGCTCTGACACAAATGTCTCTCCAAATCTCTGAAACCAACAGGTCGC 4351
Db 1354 TCCATCTACGTCTCTGGATAATAAGTAGTTTACAGATCACTGAAACCCGTACGGTCCGC 1413
Qy 4352 ATTGTCCCGGGAGGCCCATGCACTGCCAGTCCCTGGCATTGACCACTTCTGTCTAAGC 4411
Db 1414 ATCGTCCCGGGGCCCATGCACTGTCACTCTCCCTGGAGTGG--AATACCCGGTGGG 1470
Qy 4412 AAGTGGCCATCCAGCAACCTCTGGAGTCAAGCACCGCTTGGCTGTTCACACAAATGGG 4471
Db 1471 AAGCACCGGTTCAGACCAACCTCTGGAGTCAAGCACCGCCATCTGTCTCTACAGCGGG 1530
Qy 4472 GTCTGTATATTGCTGAGACTGATGAGAAAAGATCAACCGCATCAGGACAGGTCAACCA 4531
Db 1531 GTCTTTTACATCAGGAACTGATGAGAGAGATCAACCGATAGGCAGGTCAAGCA 1590
Qy 4532 AGTGGAGATCTCATCTGTTGTGGGGCCCCAGTGGCTGTGACTGTGTAATAATGATGCC 4591
Db 1591 GACGGGAGATCTCTTGTGTGGGTGGGATACCTTCGGGAATGTGACTGCAAGAACGACGCC 1650
Qy 4592 AACTGTGATGTTTTCTGGAGACGATGTTATGCCAAGGATGCAAGTAAATACCCCA 4651
Db 1651 AACTGTGATGTTTACCAAGCGGAGACGGCTAGCCAAAGATGCCAACTCAATGGCGCG 1710
Qy 4652 TCTTCTCTGCTGTGTGTGTGGGAGCTCTACGTGGCCGACCTTGGGAACATPCGA 4711
Db 1711 TCTCTCTGCGCGCTCGCCAGATGGCACTCTGTATATTGCAGATCTGGGAATATCAG 1770
Qy 4712 ATTGGGTTTACGGAGACAGCTTTCTTCAACACCCAGACATGATGAGTGTCT 4771
Db 1771 ATCGGGCGGTTTGGAAAGATAACCTTACTGAATCAATGAATTTTACGAAGTTGCC 1830
Qy 4772 TCACCAATTGACGAGGCTCTATCTGTTTGATACCAACCGGCAAGCACCTGTACACCCAA 4831
Db 1831 TCTCCACTGATCAAGAGCTCTACATCTTTTGACATCAACGGTACTCACAGTACACCGTG 1890
Qy 4832 AGCTGCCCAACAGAGACTACTGTGTAACCTTCACTTCACTGGGAGCGGCGACATCA 4891
Db 1891 AGCTGTCTACGGGTGACTACTATATAATTTTAGTTACAGCAATGACAATGACGTCAAC 1950
Qy 4892 CTCATACAGACAACATGCAATGGTAAATGTCCGCGAGCTCTACTGGGATGCC 4951
Db 1951 GCTGTAATGACCAATGCAACCCCTCCGAATCCGAAGGATCCGAATCGGATGCCG 2010
Qy 4952 CTCTGGTGTGTGTCAGATGCCAGGTTGATCTGGGTGACCATGGGCAACCAAGTGA 5011
Db 2011 GTGGGGTGTGTCTCTGATACCAAGGTGATGTTGATGATGATGATGATGATGATGATG 2070
Qy 5012 CTCAGAGTGTGACCAACAGGACAGGAGTGGCCCATGATGATGATGATGATGATGATGATG 5071
Db 2071 CTGAAAAGCAATGACCGCTCAGGCGCTGGAACTGGTTTGTGTTTACTTACCATGGCAAC 2130
Qy 5072 GGCCTTCTGGCAACCAAAAGCAATGAAACCGATGGACAACTTTTATGAGTACGACAGC 5131
Db 2131 GGGCTTTTAGCCACCAAAAGTACGAACTGGATGGACAACTTTTGTGACTATGACATG 2190
Qy 5132 TTTGGCGGCTGCAAAATGTGACTTCTCTCTGCGGCGAGGTGAGAGTTTCCGAGTGAT 5191
Db 2191 GAAGGTGCGCTGAGGAATGTTACTTCTCCCACTGGGGTGGTTTACAAAACCTGCAACGGGAC 2250

Qy 5192 ACAGACAGTTTCAGTGCATGTCCAGGTAGAGACCTCCAGCA--AGGATGATGTCAACCAT 5248
Db 2251 ATGGACAAGGCTATACAGGTGGACATCGAGTCACTCCAGCAGAGGAAGATGTGAGTCC 2310
Qy 5249 ACCACCAACCTGTGTGCTCAGGCGCTTTTACACACTGTCTGCAAGACCAAGTCCGGAAC 5308
Db 2311 ACTTCGAATTTGTCTCATCGATTCTCTACACCATGTCTCAAGACCAAGTAAAGAAC 2370
Qy 5309 AGCTACTATATCGGGGGCGATGGCTCTTGGGGCTGTCTGTGCGCAACGATGGAGGTG 5368
Db 2371 AGTTACAGATTTGGGTATGATGGCTCCCTTAGAATCTTCTATGCCAGTGTCTGGACTCT 2430
Qy 5369 GCGCTGCAGACTGAGCCCACTTCTGTGCTGGCAACCGCTCAACCCCAACCGTGGGCAAGAG 5428
Db 2431 CACTACAGACAGACCCACAGTTCTGGCTGGCAACCGGAATCCACAGTAGCCAAAGA 2490
Qy 5429 AATGTACGCTGCCCATCGAACACCGCTCAACCTGTGTGAGTGGCCCGCAGCGCAAGAG 5488
Db 2491 AACATGACTTCTCCGGTGGAAACCGGCAAGATCTGTGGAGTGGAGTTCGGAAGAA 2550
Qy 5489 CAGCTCGGGCCAGGTCACTGTCTTTGGGCGCGCTGGGGTGCACAAACCGAAATCTC 5548
Db 2551 CAAGCCAGGCAAAAGTCAAGTATTCGGCGGAGCTCAGGTCATATGGGCGCAACTA 2610
Qy 5549 CTATCTCTGACTTTTGTATCGCTTAACACGACAGAGAAAGATCTATGATGATGATGATG 5608
Db 2611 CTCTCAGTGGACTTTTGTATCGGACCAACCAAGACGGAAGATCTATGATGATGATGATG 2670
Qy 5609 TTCACCTTTCGATTTCTAGCACAGCGGGCGGCCAGCTCTGCTCAGCCAGCAGC 5668
Db 2671 TTTCTCTGAGGATCGGTTACGACACGTCGGGGCAACCGACTCTCTGCTGCGGAGTAGC 2730
Qy 5669 AGGCTGAATGTGTCAACGTGACATCTCCCTGGGGGTTACATGCTGGCATCCAGAGG 5728
Db 2731 AAGCTATGCGAGTGAACGTCACTCTATCCACCGTCAAAATGCGCAGATCCAGAGA 2790
Qy 5729 GGCATCATGTCTGAAAGATGGAATACGACAGCGGGCGGCATCATCATCAGATCTTC 5788
Db 2791 GGGACACAGACGAAAGTGGACTATGACAGCCAGGGGAGGATCGTATCTCGGCTTTT 2850
Qy 5789 GCTATGGAAGACATGAGGTACACATCTTAGAGAAAGTCCATGCTGCTACTACAC 5848
Db 2851 GCCATGGGAAACATGAGGTACACGTAATTGGAAGTCCATGTTCTTCTGCTCCAT 2910
Qy 5849 AGCCAGAGGCAATATCTTTGAGTTCGAAAGATGACCGCTCTCTTCTGTGACGATG 5908
Db 2911 AGCCAGGCGCAGTACATCTTCGAATAGCATGTGGGACCGCTGTCCGCAATCACCATG 2970
Qy 5909 CCCAACGTGGCGCGGACGACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATC 5968
Db 2971 CCCAGTGTGGTTCGCCACACCATGACACCATCCGCTCCATTTGGCTACTACCGCAATC 3030
Qy 5969 TATCAGCCCCCTGAGGGCAATGCTCAGTCAATACAGGACTTTCAGTGGATGGGACCTC 6028
Db 3031 TACAATCCCCCAGAAAGCAATGCTCTATCATCACCGACTACACAGGAGGAGGCTGCT 3090
Qy 6029 CTTACACCTTCTACTTGGGCACTGGCGGCGAGGTTGATATACAGTATGCAAACTGTC 6088
Db 3091 CTGCAAAACAGCTTCTCTGGGAAACGAGTCGGAGGGTCTTATTTCAAGTATAGAAAGCAG 3150
Qy 6089 AAGCTGGCAGAGACGCTCTATGACCAACCAAGGTGAGTTTTCCTATGACGAGACGGCA 6148
Db 3151 AGGCTATCAGAAATTTTATACGACAGCAAGAGTCAAGTTCCTTACCTACGACGAAACAG 3210
Qy 6149 GGCATGTGGAAGACATCAACCTACAGATGAGGGCTTTCAGTGCACCATCCCGTACCGT 6208
Db 3211 GGAGTCTCTGAAACAGTAAACCTTCAGAGTGTGGTTTTTATTTGCAACCATTAGATACAG 3270
Qy 6209 CAGATTGGGCGCTGATTGTCGCGGACAGATCTTCCGCTTCTACAGGAGGATGCTGCTCAAC 6268
Db 3271 CAATTTGTTCCCTGATTTGACAGACATTTTCCGCTTCAGCGAGGATGGAATGGTAAAT 3330

Qy 6269 GCCGGTTTGGCTACAACTATGACCAACAGCTTCGGGTGACAGGATGAGGCTGTGATC 6328
Db 3331 GCGAGATTTTGGCTATAGTACGACCAACAGCTTCGGGTGACAGGATGAGGCTGTGATC 3390
Qy 6329 AACGAGACCCCACTGCCATTTGATCTCTATCGCTATGATGATGTGAGGCAAGACAG 6388
Db 3391 AATGAACACCACTGCCATTTGATCTCTATCGCTATGATGATGTGAGGCAAGACAG 3450
Qy 6389 CAGTTTGGGAAGTTTGGTGTCACTTACTATGACATTAACAGATCATACCAAGCTGTG 6448
Db 3451 CAGTTTGGGAAGTTTGGGAGTGATATCTACGACATCAACCAAAATCATTTCCACGGCGTG 3510
Qy 6449 ATGACCCACACCAAGCACTTTGATGATGATGATGATGATGATGATGATGATGATGATG 6508
Db 3511 ATGACTTATACAAAGCACTTTGATGATGATGATGATGATGATGATGATGATGATGATG 3570
Qy 6509 TTCCGCTCGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6568
Db 3571 TTTAGGTCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3630
Qy 6569 AAGGAGCTGAAGTAGGACCTTACGCAATACCACTCGCTACTCTCTATGATGATGATGATG 6628
Db 3631 AGAGAGATTAAATTTGGGCTTTTGGCAACACATCAACAAATACCGGTCAGGATGACGCTG 3690
Qy 6629 GAGCGGAGCTGACAGACAGCTCTCCATCAATGACAAAGCACTCTGGGCTACAGCTACGAC 6688
Db 3691 GATGGACAGCTCCAAACAGTTTACCTAAACGAAAGATCATGTGGCGGTACAACTACGAC 3750
Qy 6689 CTCATGCGGAACCTGACCTTACTAGGCGCTGGGAAAGTGCAGCGGTCAACACCACTACGG 6748
Db 3751 CTAATGGAAACCTTCCACTTGTCTCAACCCGAGCAGCGCGGCTGACCCCTCTGCGC 3810
Qy 6749 TATGACATCCGACCCCACTACCTCGCTGGGTGACGTGCAATACAAAGATGATGATGATGATG 6808
Db 3811 TATGACCTCGGACAGCAATACCCGCTGGGCGATGTTCACTACCGGCTGGATGATGATGAT 3870
Qy 6809 GGCTTCCTGAGGACCGGGCGGTGATATTTTGGTACAACTACAGTGGCTGCTGCTATC 6868
Db 3871 GGTTTCCTGCTGACAGGGGCACTGAAATTTTGAATACAGCTCCAAAGGCTTCTGACT 3930
Qy 6869 AAGGCTTACACCGGCTGCGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 6928
Db 3931 CGAGCTTACAGTAAGGCACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3990
Qy 6929 CGCGTGTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 6988
Db 3991 CGTGTCTTACGAAACACAGCTGGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4050
Qy 6989 AACCCCAAGGTCACCCAGCTGTACCAACCACTCCAGCTCTGAGATCACTCCCTCTAC 7048
Db 4051 TACCCCAAGGATTTACTCAGCTACCAACCACTCCAGCTTCCAGTTCAGAAATCACTCCCTGTAC 4110
Qy 7049 TACGACTTGCAAGGACACCTCTTTGGCATGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 7108
Db 4111 TATGACCTCCAGGACATCTTTCCGATGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4170
Qy 7109 GCTTGTGACAACTCGGAGCCCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7168
Db 4171 GCCTCGGACAAACAGCGGACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4230
Qy 7169 CAATCTGTACACAGCTATGGGAGATCTACATGGATACCAACCCCACTTCCAGATC 7228
Db 4231 CAGATCAGTACATGCTGCTATGGTGGATCTACTTTGATCTCCACAGTCGACTTTCAGCTG 4290
Qy 7229 ATCATAGCTACCTATGGGCTCTATGATCCACTCAACCAAGTTTGTCCACATGGGCGG 7288
Db 4291 GTAATTGATTTCCAGCGGGCTTGTATGACCCGCTCAACCAACTATCCACTTTGGAGAA 4350
Qy 7289 CGAGATTATGATGCTGGCGGAGCTGTGACTGACCCAGACACAGCTGTGGAAGCAC 7348
Db 4351 AGAGATTATGACATTTTGGCGGAGATGGACCAACCGGACATTTGAATCTCGA---AA 4407
Qy 7349 CTTAGTAGCAGCAACGTCATGCTTTTAAATCTCTATATGTTCAAAAAACAACACCCCATC 7408

Db 4408 AGGATCGGAAAGGACCCCTGCTCTTTTAACTGTATATGTTTTCGGAATAACAACCCCGC 4467
Qy 7409 AGCAACTCCAGGACATCAAGTGTCTTTCATGACAGATGTAACAGCTGCTGCTCCTTT 7468
Db 4468 AGCAAAATCCATGATGTGAAAGATTACATCAGGATGTAACAGCTGCTGCTGCTGCTGCT 4527
Qy 7469 GGATTCCTAGCTACACAACTGATCCCTGGTTATCCCAACAGACATGGATGCTGAA 7528
Db 4528 GGCTTCCATCTGCAATGCTATTCCTGGATTCCTGTTCCCAATTTGATTTAACTGAG 4587
Qy 7529 CCCTCTACGAGCTCATCCACACACA 7554
Db 4588 CTTTCTATGAGCTTGTGAAGATCA 4613

RESULT 7
AK037897 3038 bp mRNA linear HTC 19-SEP-2003
LOCUS
DEFINITION Mus musculus 16 days neonate thymus cDNA, RIKEN full-length
enriched library, Clone: A13006L24 product: odd Oz/ten-m homolog 1
(Drosophila), full insert sequence.
ACCESSION AK037897 GI:26332325
VERSION
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20493374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3038)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirakawa, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Kawai, J., Kojima, Y., Kondo, H., Konno, H., Kouda, M.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, K., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/.

FEATURES

Location/Qualifiers
1..3038
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A130606L24"
/db_xref="MGI:2402337"
/db_xref="taxon:10090"
/clone="A130606L24"
/tissue="thymus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days neonate"
1..2978
/note="unnamed protein product; odd Oz/ten-m homolog 1 (Drosophila) (MGI:MGI:1345185, GB|NM_011855, evidence: BLASTN, 99%, match=3037)
putative"
/codon_start=3
/protein_id="BAC29893.1"
/db_xref="GI:26332326"
/translations="EINLSSEPHILAGAVNTLKGKNSLIPGHNANLIEWRQREKVN KGNVAFERLRANLNLISIDPDHTRGKIYDDHRTKTLRLIYDQTPRLWSPVS RYNEVALTSPGLVFIORGWNEKWEYDQSKLISRTWADGKINSYILEKSMLL LHSQRYIFPYDQSDCLLSTWFSWVHSLQMLSVGYRNYTPDSSTFIQDYSR CTIRYRTGLIQRQIFRSEBQVNLVNFYDQSDCLLSTWFSWVHSLQMLSVGYRNYTPDSSTFIQDYSR DVSGRTQFGKFIYDQSDCLLSTWFSWVHSLQMLSVGYRNYTPDSSTFIQDYSR YDNNRMTIPIYDQSDCLLSTWFSWVHSLQMLSVGYRNYTPDSSTFIQDYSR HGNSARLPIYDQSDCLLSTWFSWVHSLQMLSVGYRNYTPDSSTFIQDYSR GWTVQYVYDGLGRVASKSLQHLQFFVADLANPRTVTHLYNHTSAEITSLYDLQGL HLAMELSSEEEYVACDNGTIFVAFSRGQVKEILYTPGDIHPDTPDFVILQ FHGLYDFTLQVHLQGLDYVAVGWTPNNHHIWKQNLKPLNLYSFENNYVVGK IQVAYCTDITGLFWLQFNLVLPFGPKPELENMELTYELLQITQWDPGKM ILGQICQLQKLNFI.SLQLPMTQYNEGRCLLEGQKPRFAVPSVFKGKFKAIKE GIVTADIIGVANDSRLLAAILNNAHLYNHLFTIBGRDTHYFVILGSLLEEDLVNIGN TGRRRLIENGWVTVSQMVSVLNLTFRFPADLOLQHGALCFNIRVGTVEEKNHVL MARQRAVAQWQEQRLQEGEBGTRVTEGEKQQLLGTGRVQGVDFVLSVEQYLE LQSDANNHMPHROSEIGRR"

CDS

1..2978
/note="unnamed protein product; odd Oz/ten-m homolog 1 (Drosophila) (MGI:MGI:1345185, GB|NM_011855, evidence: BLASTN, 99%, match=3037)
putative"
/codon_start=3
/protein_id="BAC29893.1"
/db_xref="GI:26332326"

ORIGIN

Query Match 16.9%; Score 1407.8; DB 11; Length 3038;
Best Local Similarity 67.4%; Pred. No. 8.9e-276;
Matches 2014; Conservative 0; Mismatches 967; Indels 6; Gaps 2;
QY 5362 GGAGTGGCCCTGAGACTGAGCCGACCTTGCTGGCTGGCAGCGTCAACCCACCGTGGG 5421
DB 2 GGAGATCAACTGCTCGAGAGCCGACATCTCTGGCAGGGGCTGTCAACCCACCTTAGG 61
QY 5422 CAAGAGGAATGTCACTGCTGCCATGCAACACGCGCTCAACCTGGTGGAGTGGCGCCAGG 5481

62 CAAATGCAACATATCGTTTCTGGAGAGCACAATGCAAACTATTGAGTGGAGACAGAG 121
QY 5482 CAAAGAGCAGGCTCGGGCCAGGTCACCTGTCTTTGGGGCGCGGTGGGGTGCACAAACCG 5541
DB 122 AAAGAGGCAAAACAAGCAACAGTTTTCAGCTTTTGGAGAGAGACTGAGGGCCCAATACAG 181
QY 5542 AAATCTCTATCTCTGGACTTTGATCGGTAAACAGCAGCAGAGAGAGATCTATGATGACCA 5601
DB 182 AAACCTACTTTTCCATTTGATTTGATCATATGACCGCAGCGGTAAAGTCTATGATGACCA 241
QY 5602 CGCAGAGTTTCAACCTTCTGGATTCTGACAGCAGCGGGGGCGGCGGCTCTGCTGCTCACC 5661
DB 242 CGAAATTTCACTCTTCTGAAATCTTTTATGACAGCAGCGGGGACCTATTCTGTGCTCTCC 301
QY 5662 CAGCAGCAGGCTGAATGGTGTCAACGCTGACATCTCCCTTGGGGGTTTACATTTGCTGGCAT 5721
DB 302 TGTGAGCAGATATAATGAAGTGAACATCACGTTATTCACCTTCAGGATTAGTCACATTTAT 361
QY 5722 CAGAGGGGATCATGCTCTGAAAGATGGAATAGCAGCAGCGGGCGGCGCATCACATCCAG 5781
DB 362 TCAAAGAGGAACTCGAATGAAAGATGGAGTATGACGAGAGTGGAAAAATAATCTCAAG 421
QY 5782 GATCTTCTGCTGATGGGAAGACATGAGCTACACATCTTACAGAGTCCATGGTGGTCTGCT 5841
DB 422 AACTTGGGCCGATGGGAAATTTGGAGCTATACCTTCTAGAAAGTCTGTGATGCTTCT 481
QY 5842 ACTACAGCAGCAGGAGTATATCTTTGAGTTCGAAAGATGACCGCTCTCTTCTCTGCT 5901
DB 482 GTTACACAGTCAAGCGGCTTACATCTTTGAGTATGACCAATCAGATTCGCTGTGCTCAGT 541
QY 5902 GACGATGCCAACGTTGGCGGCGAGACACATAGACACCATCGCTCAGTGGGCTACTACAG 5961
DB 542 TACATGCTAGATGTTGGCCACAGCTTACAAACCATGTTTTCAGTGGGCTACTACCG 601
QY 5962 AAACATCTATCAGCCCTTGGGCAATGCTCAGTACATACAGGACTTCTACTGAGGATGG 6021
DB 602 TAATATCTACACCCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 661
QY 6022 GCACCTCTTACACCTTCTACCTGGGCTGGGCGGAGGTATATACAGATGACCA 6081
DB 662 CCGAGTGTACAGACTCTGATCTAGGAGTGGGCGCAGAGTTTTATACAGTACACAAA 721
QY 6082 ACTGTCAAAGCTGGCAGAGAGCGCTCTATGACACCAACCAAGCTCAGTTTCCCTATGACGA 6141
DB 722 CAAAGCAAGCTGTCTGAGATTTCTATGACACCACTCAGTCACTAATACATATGAAGA 781
QY 6142 GACGGCAGGATGCTGAAGACCATCACTTACAGATGAGGGCTTCACTTGCACCATCCG 6201
DB 782 GTCTTCTGGAGTAAATTAAGCAATACACCTGATGATGACGGGATTCATCTGCACAATCAG 841
QY 6202 CTACCGTCAAGATTCGGGCGCTGATTGACCGACAGATCTTCCGCTTCACTCAGGAAGCAT 6261
DB 842 ATACAGCAACACAGACCTCTCATGACGCTCAGATTTTCGATTCACTGAGAGAGTCT 901
QY 6262 GGTCAACCGCGTTTGAATCAACTATGACAGCTTCCGGGTGACGAGCATGACGAGC 6321
DB 902 TGTGAATGACGGTTTGACTACA---GTTACAAACAACTTTTCGTTGTCACCAAGCATGCAAGC 958
QY 6322 TGTGATCAAGAGACCCCACTGCCCCATGATCTATCGCTATGATGATGATGATGATGATGATGAT 6381
DB 959 TGTGATCAAGAGACCCCTTTGCCCATAGATCTGTACCCGATATGTTGATGTTTCTGTTAG 1018
QY 6382 GACAGACAGTTTGGGAAGTTTGGTGTCTATTTACTATGACATTAACCCAGATCATCACAC 6441
DB 1019 AACAGAACAGTTTGGAAAAATTCAGTGTAAATTAATATGATTTAAATCAGGTTATAACTAC 1078
QY 6442 AGCTGTATGACCCACACCAAGCATTTTGTATGATATGAGGAGTGAAGAGTGAAGTGA 6501
DB 1079 AACAGTATGAAGCACACCAAGATTTTCAATGCAATGGACAAAGTCAATGAGGTCCAATA 1138
QY 6502 TGAGATCTTCCGCTCGCTCATGTACTGATGAGCGCTCAGTATGATAAATGAGGTGGGAGT 6561
DB 1139 TGAGATCTTAAAGCAATTCGCTACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS

Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akizawa, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, F., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanishi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1751-1771 (2000)
20530913
11076961
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 1284)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuura, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnishi, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tonari, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. .1284
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:A330048C04"
/db_xref="MGI:2403942"
/db_xref="taxon:10090"
/clone="A330048C04"
/sex="male"
/tissue_type="spinal cord"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1. .1284
misc_feature
/notes="neuregulin 1 (MGD|MGI:96083, MGD|MGI:1345182, GB|NM_011858, evidence: BLASTN, 99%, match=186)"
ORIGIN

Query Match	8.9%	Score 743.8	DB 11	Length 1284
Best Local Similarity	86.3%	Pred. No. 1.7e-140		
Matches 845	Conservative 0	Mismatches 107	Indels 27	Gaps 1
QY	2195	GGACACGAGCTGTTATCGAGATCTGTGCGGACTGTGGTGGCCATGGGTGGCTA	2254	
DB	80	GGACAAACCCAAAACCCAGATCTGTGCGGCTGACCTGTGGTGGCCATGGGTGGCTA	139	
QY	2255	GCGGCGACCTCCCGCTCGAGGATGGGTGGATGGGGGCGAGCTCGACACGCGGCTGC	2314	
DB	140	GGAGGCACCTCCCGCTGTGAGGATGGTGGATGGGGGCGGATCGACACGCGGCTGC	199	
QY	2315	CACCGGCGCTGTGCGGAGATGGGACCTGCGGCGACGGCAAGTCGAGTCGAGCCCTGC	2374	
DB	200	CACCCACGCTGTGCAGAACACGGGACCTGCGGCGACGGCAAGTCGGAATCGACCCGCGC	259	
QY	2375	TGGAATGCGCAACACTGCACCATCGCTACTATCTGATAGGTAGTGTAAAGAGGGTTGC	2434	
DB	260	TGGAATGCGAGCACTGCACCATC-----GAGGCTGT	292	
QY	2435	CCTGGTTGTGCAATGGCAACGGCAGATGACCTTAGACCTGAAATGGTGGCTGCGT	2494	
DB	293	CCTGGCTTGTGCAATGGAAATGGCAGATGACCTGGACCTGAAATGGTGGCTGCTGTC	352	
QY	2495	TGCCAGCTGGGCTGGAGAGGAGCTGGCTGTGACACTTCCATGGAGAGCTGCTGGGTGAC	2554	
DB	353	TGCCAGCTGGGCTGGAGAGGAGCTGGCTGGACACATCCATGGAACGGGCTGTGGAGAC	412	
QY	2555	AGCAAGACCAATGATGGAGATGGGCTGGTGGACTGCATGGACCTGACTGCTGCTCCAG	2614	
DB	413	GGCAAGACCAACGACGAGATGGCTTGGTGGACTGCATGGACCTGACTGCTGCTCCAG	472	
QY	2615	CCCTGTGCCATATCAACCGCTGTGCTGGCTCCCTAACCTTGGACATCATCCAG	2674	
DB	473	CCTCTGTGTCATGTCACACCGCTGTGCTAGGCTCCCTGACCCCTGGACATCATCCAA	532	
QY	2675	GAGACACAGGTCCTGTGTGCACAGCAACCTACACTCTCTTATGACCCGATCAAGTTC	2734	
DB	533	GAGACACAGGTCCTGTGTATCCAGCAACCTGACACTCTTCTACGATCAATCAAGTTC	592	
QY	2735	CTCGTGGCGAGGACAGCAGCAGCATATCCCGGGAGAACCCCTTTGATGAGGGCAT	2794	
DB	593	CTCGTGGCGAGGACAGCAGCAGCATATCCCGGGAGAACCCCTTTGACGGAGGCGAT	652	
QY	2795	GCTTGTGTATTCTGTGCGCAAGTGCATCATCATGAGTGAACCCCTGCTGTTGTGAAC	2854	
DB	653	GCTTGTGTATTCTGTGCGCAAGTGCATCATCATGAGTGAACCCCTGCTGTTGTGAAC	712	
QY	2855	ATCAGTTTTGTCAATAACCTCTCTTTGGATATACATCAGCAGGCAAGATGCGCTTT	2914	
DB	713	ATCAGTTTTGTCAATAACCTCTCTTTGGATATACATCAGCAGGCAAGATGCGCTTT	772	
QY	2915	GACCTGTGTAATGCGGCACTCTCCATCATCTCGGTTCCGAGCGGGCACCTTTCATC	2974	
DB	773	GACCTGTGTAATGCGGCACTCTCCATCATCTTTGAGATTCGAAAGGACCCCTTTCATC	832	
QY	2975	ACACAGGAGCACACCTCTGTGGCTGCCATGGGATGCGCTTCTTTGATGAAACCATCATC	3034	
DB	833	ACACAGGAGCATACCTCTGTGCTACCTGGATGCGCTTCTTTGATGAAACCATCATC	892	
QY	3035	ATGAGACATGAGGAGATGAGATCCAGCTGTGACCTGAGCAATTTTGGCCGCCCAAC	3094	
DB	893	ATGAGACATGAGGAGATGAGATCCAGCTGTGACCTGAGCAATTTTGGCCGCCCAAC	952	
QY	3095	CCAGTGTCTCTCCATCCCACTGACGTCTCTCGCAGCTCTCTGTGAGAGAAAGGCCCC	3154	
DB	953	CCTGTGTCTCTCCATCCCACTGACATCTTTCGCCAGCTCTCTGTGAGAGAAAGGCCCC	1012	
QY	3155	ATTGTGCGGAAATTCAGG 3173		
DB	1013	ATTGTGCGGAAATTCAGG 1031		

RESULT 9	AL045768	DKFZp434P206.r1	434 (synonym: htes3)	Homo sapiens	cDNA clone	EST 04-SEP-2003
LOCUS	AL045768	DKFZp434P206.5	5	mRNA sequence		
DEFINITION	AL045768	EST	1	GI:54333880		
ACCESSION	AL045768	EST	1	GI:54333880		
VERSION	AL045768	EST	1	GI:54333880		
KEYWORDS	EST	1	GI:54333880			
SOURCE	INGOLSTAEDTER Landstr.1, D-85764 Neuherberg, Germany					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 718)					
AUTHORS	Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.					
TITLE	EST (Koehler, et al.)					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: MIPS					
FEATURES	1..718					
source	Location/Qualifiers					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/clone="DKFZp434P206"					
	/tissue_type="testis"					
	/dev_stage="adult"					
	/lab_host="DH10B"					
	/clone_lib="434 (synonym: htes3)"					
	/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"					
ORIGIN						
Query Match	8.5%; Score 713.8; DB 9; Length 718;					
Best Local Similarity	99.7%; Pred. No. 1.6e-134;					
Matches	715; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
QY	7431	GCCTCATGACAGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAAAGTGA	7490			
Db	1	GCCTCATGACAGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAAAGTGA	60			
QY	7491	TCCCTGGTTATCCCAACCAAGACATGGATGCCATGGAAACCTCTCTAGAGCTCATCCACA	7550			
Db	61	TCCCTGGTTATCCCAACCAAGACATGGATGCCATGGAAACCTCTCTAGAGCTCATCCACA	120			
QY	7551	CACAGATGAAACCGAGGTGGGACACACAGCAAGTCTATCTCCGGGTACAGTGTCAAG	7610			
Db	121	CACAGATGAAACCGAGGTGGGACACACAGCAAGTCTATCTCCGGGTACAGTGTCAAG	180			
QY	7611	TACAGAGCAGCTCAAGGCTTTGTACCTTAGAACCGTTTGACCACTCTATGGGTCCA	7670			
Db	181	TACAGAGCAGCTCAAGGCTTTGTACCTTAGAACCGTTTGACCACTCTATGGGTCCA	240			
QY	7671	CAATCAGCAGCTGCACAGGCTCCAAAGACCAAGAGTTTGCATCCAGGGCTCAGTCT	7730			
Db	241	CAATCAGCAGCTGCACAGGCTCCAAAGACCAAGAGTTTGCATCCAGGGCTCAGTCT	300			
QY	7731	TTGCAAGGGGGTCAAGTTTCCCTTGAAGGATGCCGAGTGCCAGAGATCATCATGCT	7790			
Db	301	TTGCAAGGGGGTCAAGTTTCCCTTGAAGGATGCCGAGTGCCAGAGATCATCATGCT	360			
QY	7791	TGGCCAATAGGATGGGCGAAGGTTGCTGCCAATCTTTGAACCAATGCCCATACCTAGAGA	7850			
Db						

Db	361	TGGCCAATAGGATGGGCGAAAGGGTTGCTGCCATCTTTGAACCATGCCCATCTACCTAGAGA	420			
QY	7851	ACCTGCACCTTACCAATTTGATGGGTGGATACCCATTACTTTTGTGAAACCGAGACCTTCAG	7910			
Db	421	ACCTGCACCTTACCAATTTGATGGGTGGATACCCATTACTTTTGTGAAACCGAGACCTTCAG	480			
QY	7911	AAGGTGACCTGGCCATCTCTGGGCGCTCAGTGGGGGGCGGCGAACCCTTGAGAAATGGGGTCA	7970			
Db	481	AAGGTGACCTGGCCATCTCTGGGCGCTCAGTGGGGGGCGGCGAACCCTTGAGAAATGGGGTCA	540			
QY	7971	ACGTCACTGTGTCCAGATCAACACAGTACTTAATGGCAGGACTAGAGCTACACAGACA	8030			
Db	541	ACGTCACTGTGTCCAGATCAACACAGTACTTAATGGCAGGACTAGAGCTACACAGACA	600			
QY	8031	TCCAGCTCCAGTACGGGCGACTGTCTTGAACACACACAGCTACGGCAACAGTTGGATGAGG	8090			
Db	601	TCCAGCTCCAGTACGGGCGACTGTCTTGAACACACACAGCTACGGCAACAGTTGGATGAGG	660			
QY	8091	AGAAGGCACGGGTCTCTGGAGTGGCCCGCGAGAGCGCTGCGCCCAAGCGTGGGCCC	8147			
Db	661	AGAAGGCACGGGTCTCTGGAGTGGCCCGCGAGAGCGCTGCGCCCAAGCGTGGGCCC	717			
RESULT 10	CF534506	856 bp	mRNA	linear	EST 12-SEP-2003	
LOCUS	CF534506	IMAGE:30536758.5	5	mRNA sequence		
DEFINITION	CF534506	EST	1	GI:34586474		
ACCESSION	CF534506	EST	1	GI:34586474		
VERSION	CF534506	EST	1	GI:34586474		
KEYWORDS	EST					
SOURCE	Mus musculus (house mouse)					
ORGANISM	Mus musculus					
REFERENCE	1 (bases 1 to 856)					
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/					
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)					
JOURNAL	Unpublished (1999)					
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@bbs-r@mail.nih.gov Tissue Procurement: Dr. Jim Lin, University of Iowa cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa DNA Sequencing by: Dr. M. Bento Soares, University of Iowa Clone Distribution: Distribution information can be found at http://genome.uiowa.edu/distribution/mousefl.html This clone was contributed by the Brain Molecular Anatomy Project (BMAP)					
Seq primer:	pYX-5					
Location/Qualifiers	1..856					
	/organism="Mus musculus"					
	/mol_type="mRNA"					
	/strain="C57BL/6"					
	/db_xref="taxon:10090"					
	/clone="IMAGE:30536758"					
	/tissue_type="whole brain"					
	/dev_stage="embryo 13.5, 14.5, 16.5, 17.5 dpc"					
	/lab_host="DH10B (T1 phage resistant)"					
	/clone_lib="NIH BMAP G10"					
	/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is ACGGACAGAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the					

QY 6744 TACGGTATGATCCCGGACCGGATCACTCGGCTGGGTGAGTGCATACAGATGGATG 6803
 Db 241 TACGGTATGATCCCGGACCGGATCACTCGGCTGGGTGAGTGCATACAGATGGATG 300
 QY 6804 AGGATGCTTCTTCTGAGGACGCGGGGCGGTGATATCTTTGAGTACAACTCAGCTGGGCTGC 6863
 Db 301 AGGATGCTTCTTCTGAGGACGCGGGGCGGTGATATCTTCTGAGTACAACTCAGCTGGGCTGC 360
 QY 6864 TCATCAAGGCTCAACACCGGCTGGCAGCTGAGTGCAGTACCGCTACGATGGGCTGG 6923
 Db 361 TCATCAAGGCTCAACACCGGCTGGTGGTGGAGTGTGAGTACCGCTACGATGGGCTGG 420
 QY 6924 GGGGGCGGTGTCAGCAAGAGCAGCCACAGCCACACCTGCGAGTTCTTCTATGCAAGACC 6983
 Db 421 GACGCCAGTATCCAGCAAGAGCAGCCACAGCCACACCTGCGAGTTCTTCTATGCAAGACC 480
 QY 6984 TGACCAACCCCAACAGGTCAACCACTGACCACTGACCACTCCAGCTCTGAGATCACCTGCC 7043
 Db 481 TGACCAACCCCAACAGGTCAACCACTGACCACTGACCACTCCAGCTCTGAGATCACCTGCC 540
 QY 7044 TCTACTAGGCTTCAAGGACACCTCTTGGCATGGAGCTGAGCTGGTGTGATGATTTT 7103
 Db 541 TCTACTATGATCTCAAGGACACCTCTTGGCATGGAGCTGAGCTGGTGTGATGATTTT 600
 QY 7104 ACATAGCTTGTGACCAATCGGACCCCTCTTGTCTTTAGTGAACAGGTTTGTATGA 7163
 Db 601 ACATAGCTTGTGACCAATCGGACCCCTCTTGTCTTTAGTGAACAGGTTTGTATGA 660
 QY 7164 TCAAGCAATCCTGTAACAGCTTATGGGAGATCTACATGGATACCAACCCCACTTTC 7223
 Db 661 TTAAGCAGATCTTATACAGCGGCTATGGGAGATCTACATGGATACCAACCCCACTTTC 719
 QY 7224 AGATCATATAGGCTTACCATGCTGGCTCTATGATCCACTCACCAGCTTGTCCACATGG 7283
 Db 720 AGATCATATAGGCTTACCATGCTGGCTCTATGATCCACTCACCAGCTTGTCCACATGG 775
 QY 7284 GCCGGGAGATATGATGTGCTGGCGGACGTGGATGACC 7325
 Db 776 GCCGACGGGA-TATGATGTGCTGGCTGG-CGCTGACCAAGCC 815

RESULT 12
 BG036207 870 bp mRNA linear EST 24-JAN-2001
 LOCUS 602326960F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4428351 5',
 DEFINITION mRNA sequence.

ACCESSION BG036207
 VERSION BG036207.1 GI:12431132
 KEYWORDS EST.
 SOURCE Homo sapiens
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/BTP
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10178 row: h column: 16
 High quality sequence stop: 713.
 Location/Qualifiers
 1..870
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

FEATURES
 source

/clone="IMAGE:4428351"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_91"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 7.5%; Score 629; DB 10; Length 870;
 Best Local Similarity 98.2%; Pred. No. 3.7e-117;
 Matches 700; Conservative 0; Mismatches 5; Indels 8; Gaps 6;
 QY 904 AGGACCTCCCGCTCTTCTGACACATACACAGGTACCCAGGTACCCAGTCCAGCACAGT 963
 Db 1 AGGACCTCCCGCTCTTCTGACACATACACAGGTACCCAGGTACCCAGTCCAGCACAGT 59
 QY 964 GTACTCTCTCCGCGCGGACCCCTGCGCGGACGACCTTGGCGCGGCGGCTTTAACT 1023
 Db 60 GTACTCTCTCCGCGCGGACCCCTGCGCGGACGACCTTGGCGCGGCGGCTTTAACT 119
 QY 1024 CAAGAAGCCCTCCAAAGTACTGTAACTGGAAGTGGCAGCCCTGAGCGCATCGTCACTC 1083
 Db 120 CAAGAAGCCCTCCAAAGTACTGTAACTGGAAGTGGCAGCCCTGAGCGCATCGTCACTC 179
 QY 1084 AGCCACTCTGTCATCTCTGCTGGCATCTTGTGGCCATGACCTGTTGGCTTAACTG 1143
 Db 180 AGCCACTCTGTCATCTCTGCTGGCATCTTGTGGCCATGACCTGTTGGCTTAACTG 239
 QY 1144 GCACCTGCAGCCGATGGAGGGGACAGTATGAGATCAGGAGGACACAGCCAGCAGTTG 1203
 Db 240 GCACCTGCAGCCGATGGAGGGGACAGTATGAGATCAGGAGGACACAGCCAGCAGTTG 299
 QY 1204 GCTGTGCCACGACGCTCTCCCTATACCCCTCAGGGGACCTGGTTAGAGACCCCTGA 1263
 Db 300 GCTGTGCCACGACGCTCTCCCTATACCCCTCAGGGGACCTGGTTAGAGACCCCTGA 359
 QY 1264 CAGGAAAGGCAAGGAACACACAGAAAGGAAGCCAGTAGTCTTCTTCCAGAGGACAGTTT 1323
 Db 360 CAGGAAAGGCAAGGAACACACAGAAAGGAAGCCAGTAGTCTTCTTCCAGAGGACAGTTT 419
 QY 1324 CATAGATTCTGGAGAAATTTGATGTGGAGGCGAGCCCTCCAGAGAGATTCCTCTGGCAC 1383
 Db 420 CATAGATTCTGGAGAAATTTGATGTGGAGGCGAGCTTCCAGAGAGATTCCTCTGGCAC 479
 QY 1384 TTTCTGGAGATCTCAAGTCTTCATAGACCATCTCTGTCATCTGAAATCAATGTCTCT 1443
 Db 480 TTTCTGGAGATCTCAAGTCTTCATAGACCATCTCTGTCATCTGAAATCAATGTCTCT 539
 QY 1444 GGGAAAGGCGAGCCCTTGGGCTTGGCATTTATGGCAGAAAGGCGCTCCCTCTTCAATAC 1500
 Db 540 GGGAAAGGCGAGCCCTTGGGCTTGGCATTTATGGCAGAAAGGCGCTCCCTCTTCAATAC 599
 QY 1501 ACAGTTTGAATTTGTGGAGCTCTGATGGAGGAGGCT-CCTAACCCAGAGGCGCGA 1559
 Db 600 ACAGTTTGAATTTGTGGAGCTCTGATGGAGGAGGCTCCCTAACCCAGAGGCGCGA 659
 QY 1560 GCTAGA-GGGGACCCCGCGCCAGTCT--CGGGGAACCTGTGCCCTCCCTCCAGC 1609
 Db 660 GCTAGA-GGGGAGGAGACCCCGCGCCAGTCTTGGGGGAACCTGTGCCCTCCCTCCAGC 712

RESULT 13
 CF536185
 LOCUS
 DEFINITION UI-M-G10-chh-a-16-0-UI-r1 NIH_BMAP_G10 Mus musculus cDNA clone
 IMAGE:30534831 5', mRNA sequence.
 ACCESSION CF536185
 VERSION CF536185.1 GI:34588165
 KEYWORDS
 SOURCE Mus musculus (house mouse)


```

ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        1 (bases 1 to 728)
TITLE          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
JOURNAL        NIH-MSC http://mgc.nci.nih.gov/
COMMENT        National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               Tissue Procurement: Dr. Jim Lin, University of Iowa
               cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
               cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
               DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
               Clone Distribution: Distribution information can be found at
               http://genome.uiowa.edu/distribution/mousefl.html
               This clone was contributed by the Brain Molecular Anatomy Project
               (BMAP)

FEATURES       Seq primer: PYX-5.
               Location/Qualifiers
               1..728
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="C57BL/6"
               /db_xref="taxon:10090"
               /clone="IMAGE:30534831"
               /tissue_type="whole brain"
               /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
               /lab_host="DH10B (TI phage resistant)"
               /clone_lib="NIH BMAP G10"
               /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
               Site 2: Not I; The library was constructed according
               to Bonaudo, Lennon and Soares, Genome Research, 6:791-806,
               1996. Denatured RNA was size fractionated on a 1% agarose
               gel. First strand cDNA synthesis was primed with oligo-dT
               primer containing a Not I site. Double strand cDNA was
               size selected according to mRNA size fraction, ligated
               with EcoR I adaptor, digested with NotI and then cloned
               directionally into pYX-Asc vector. The library tag
               sequence located between the Not I site and the polyA tail
               is AGCCAGACAG. This library was created for the University
               of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
               Developing Mouse Nervous System', supported by National
               Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
               program coordinator."

ORIGIN
Query Match      7.3%; Score 613.6; DB 14; Length 728;
Best Local Similarity 90.3%; Pred. No. 4.5e-114;
Matches 655; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 2762 ATCCCGGGGAGAACCCCTTTGATGAGGGCAGCTGTGTTATTCGTGCGCAAGTGATG 2821
DB 4 ATCCCGGGGAGAACCCCTTTGACGGAGGCATGCTTGTCATCTCGTGACAAAGTGATG 63
QY 2822 ACATCAGATGGAACCCCTTGGTTGGTGTGAACATCAGTTTGTCTAATAACCCCTCTCTTT 2881
DB 64 ACGTCAGATGGACCCCATTTGGTGGGTGAACATCAGTTTTCATCAATAACCCCTCTCTTT 123
QY 2882 GGATATACATCAGCAGGCAAGATGGCAGCTTTGACTTGTGTGACAAATGGCGCATCTCC 2941
DB 124 GGATATACATCAGCAGGCAAGATGGCAGCTTTGACTTGTGTGACAAATGGCGCATCTCC 183
QY 2942 ATCATCTCGGGTTCGAGCGGGCACCTTTTCATCACAGAGACACACCCCTGTGGCTGCCA 3001
DB 184 ATCATCTTGAGATTGGAAGGGCACCCCTTCATCACAGAGCATACCCCTCTGGCTACCC 243
QY 3002 TGGGATCGCTTCTTTGTCTGTAAGAACCATCATCATGAGATGAGAGATGAGATTCC 3061
DB 244 TGGGATCGCTTCTTTGTCTGTAAGAACCATCATCATGAGATGAGAGATGAGATTCC 303
QY 3062 AGCTGTGACCTGAGCAATTTTGGCCCGCCCAACCCAGTGTCTCTCCATCCCTCCACTGAGC 3121
DB 304 AGCTGTGACCTGAGCAATTTTGGCCCGCCCAACCCCTGCTCTCTCCATCCCTCCACTGAGC 363

3122 TCCTTGGCAGCTCCTGTGCAGAGAAAGCCGCCCATTTGTGCGGAAATTCAGGCTTTGCAG 3181
DB 364 TCCTTGGCAGCTCCTGTGTCTGAGAAAGCCGCCCATTTGTCCAGAAATCCAGGCCCTGCAA 423
QY 3182 GAGGAAATCTCTATCTCTGGCTGCAAGATGAGGTGAGCTACCTGAGCAGCGGACCCCT 3241
DB 424 GAGGAAATCGTCTATCGCTGGCTGCAAGATGAGGTGAGCTACCTGAGCAGCGGACCTCT 483
QY 3242 GGCTACAAATCTGCTCTGAGATCAGCTCACCCACCGGACCATCCCTTCAACCTCATG 3301
DB 484 GGCTATAAGTCTGCTCTGAGATCAGCTCACACCCGCCCATCCCTTCAACCTCATG 543
QY 3302 AAGGTGCACTCATGTTAGCGGTGAGGGCGGCCCTCTTTCAGGAAGTGTGCTGCTGAGCC 3361
DB 544 AAGGTGCCACCTCATGTTGGCAGTCGAGGTCGGCTCTTCAGGAAGTGTGCTGCTGAGCT 603
QY 3362 CCAGACCTGTCTTATTTATTTTCATTTGGGACAGACAGCTCTACACACGAGAGGTGTTT 3421
DB 604 CTTGACCTGTCTTATTTATTTTCATTTGGGACAGACAGATGCTTACACACGAGAGGTGTTT 663
QY 3422 GGGCTTTCAGAAAGCTTGTGTTTCGGTGTATGAATATGAATCTGCCAGATCTAATC 3481
DB 664 GGACTCTCAGAAAGCTTGTGTTTCGGTGTATGAATATGAATCTGCCAGATCTAATC 723
QY 3482 CTGTG 3486
DB 724 CTGTG 728

RESULT 14
CB524534
LOCUS          770 bp mRNA linear EST 09-JUL-2003
DEFINITION    UI-M-FYO-cey-b-20-0-UI.r1 NIH_BMAP_FYO Mus musculus cDNA clone
IMAGE:6843669 5', mRNA sequence.
ACCESSION     CB524534
VERSION       CB524534.1 GI:29357889
KEYWORDS      EST.
SOURCE        Mus musculus (house mouse)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
               1 (bases 1 to 770)
               NIH-MSC http://mgc.nci.nih.gov/
               National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
               Contact: Robert Strausberg, Ph.D.
               Email: cgapbs-remail.nih.gov
               Tissue Procurement: Dr. Jim Lin, University of Iowa
               cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
               cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
               DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
               Clone Distribution: Distribution information can be found at
               http://genome.uiowa.edu/distribution/mousefl.html
               This clone was contributed by the Brain Molecular Anatomy Project
               (BMAP)
               Seq primer: pYX-5.
               Location/Qualifiers
               1..770
               /organism="Mus musculus"
               /mol_type="mRNA"
               /strain="C57BL/6"
               /db_xref="taxon:10090"
               /clone="IMAGE:6843669"
               /tissue_type="whole brain"
               /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
               /lab_host="DH10B (TI phage resistant)"
               /clone_lib="NIH BMAP FYO"
               /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
               Site 2: Not I; The library was constructed according
               to Bonaudo, Lennon and Soares, Genome Research, 6:791-806,
               1996. Denatured RNA was size fractionated on a 1% agarose
               gel. First strand cDNA synthesis was primed with oligo-dT

```

primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is AGCCAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match	7.3%; Score 609.2; DB 14; Length 770;
Best Local Similarity	89.1%; Pred. No. 3.7e-113;
Matches 667; Conservative 0; Mismatches 81; Indels 1;	
1; Gaps 1;	
QY 3854	AGTCACAGTCCAGACACAATACTACCTGCGGCACAGACCCCATGATGGGCGGCTTC 3913
Db	23 AGTCACAGCCAGACACAATACTACCTGCTACAGCCCATGATGGGCGGCTTC 82
QY 3914	CTTTCTGACAGCAACAGCGCGGGTCTTTAAATCAAGTCCACTGTGTGGTGAAGGAC 3973
Db	83 CTGTCTGACACCAACAGCGCGGGTCTTCAAGTCAAGTCCACCCAGTGTGAAGGAC 142
QY 3974	CTTGTCAAGACTGTGAGTGTTCGGGACAGGTGACAGTCCCTCCCTTTGATGAC 4033
Db	143 CTGCTGAAGAACTCCGAGGTGTAGCAGGACTGGTGACAGTCCCTCCCTTTGATGAT 202
QY 4034	ACTCGCTCGGGGATGTGGGAAGGCCACAGAGACCACTCAACCAATCCAGGGGTATT 4093
Db	203 ACCCGCTCGGAGATGTGGGAAGGCCACAGAGACCACTCAACCAATCCAGGGGAATT 262
QY 4094	ACAGTGACAAAGTTTGGGCTGATCTATCTCGTGATGGCAACATGATCAGACGATCGAT 4153
Db	263 ACAGTGACAAAGTTTGGGCTCATTTATTCGTGGACGGCAACCATGATCAGACGTTGAT 322
QY 4154	CAGAATGGGATCATCTCCACCTGCTCGGCTTAATGATCTCATCATCAGCCCGCCACATC 4213
Db	323 CAAATGGATCATCTCCATCTTCTGGCTCCATGATCACTCTCGGCAGCCGCCCTC 382
QY 4214	AGCTGTGATCTGTGATGATATTTCCAGGTGAAGACTGGAGTGGCCACAGACTTAGCC 4273
Db	383 AGCTGTGATCTCGTCAATGGAGATTTCTCAGGTTCCGCTGGAGTGGCCACAGACTTAGCC 442
QY 4274	ATCAACCAATGGCACTACATTTATGCTCTGACACAAATGCTGCTCGAAATCTCT 4333
Db	443 ATCAACCAATGGCAATTTCTCTATGCTCTGACAAATGCTGCTCGAAATCTCTC 502
QY 4334	GAATAACCAAGGTGCGCATTTGTCGCGGAGGCCCATGCACTGCCAGGTCCTGGCATT 4393
Db	503 GAATAACCAAGGTACCATTTGCTGCGGAGGCCCATGCACTGCCAGGTCNTGGCATC 562
QY 4394	GACCACTTCTGTAAGCAAGGTGGCCATCCAGCAACCTCTGGAGTCAGCCACCGTTTG 4453
Db	563 GACCACTTCTGCTGAGCANGTAGCCATCCATGCGACCCCTGGAGTCAGCCACTGCTGT 622
QY 4454	GCTGTTTCACAAATGGGTCCTGTATTTGCTGAGACTATGAGAAAGATCAACCCG 4513
Db	623 GCGGTTTCACAAATGGAGTCTGTATCTGCTGAGCCAGCAGAAAGATCAATCGC 682
QY 4514	ATCAGGAGGTACCACTAGTGGAGAGATCTCACTGCTGGGCCCCCAGTGCTGT 4573
Db	683 ATCAGGAGGTCACTACAAGTGTGAGATCTCACTGCTGGTGTGCTGCTGCTGCTGCT 741
QY 4574	GACTGTAATAATGATGCAACTGTGATTG 4602
Db	742 GACTGTAATAATGATGCAACTGTGACTG 770

RESULT 15
AK050784
LOCUS 2627 bp mRNA linear HTC 20-SEP-2003
DEFINITION Mus musculus 9 days embryo whole body cDNA, RIKEN full-length

enriched library, clone:D030019H23 product:odd Oz/ten-m homolog 3 (Drosophila), full insert sequence.
ACCESSION AK050784.1 GI:26094117
VERSION AK050784.1
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
REFERENCE 2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
REFERENCE 3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076851
REFERENCE 4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
REFERENCE 5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2627)
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Atakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source

```
1. .2627
  /organism="Mus musculus"
  /mol_type="mRNA"
  /strain="C57BL/6J"
  /db_xref="FANTOM_DB:D030019H23"
  /db_xref="MGI:2418637"
  /db_xref="taxon:10090"
  /clone="D030019H23"
  /tissue types="whole body"
  /clone_lib="RIKEN full-length enriched mouse cDNA library"
  /dev_stage="9 days embryo"
1. .2627
misc_feature
  /notes="odd Oz/ten-m homolog 3 (Drosophila)
(MGI|MGI:1345183, GB|NM_011857, evidence: BLASTN, 99%,
match=1753)"
```

ORIGIN

```
Query Match          7.2%; Score 602.6; DB 11; Length 2627;
Best Local Similarity 64.3%; Pred. No. 1.8e-111;
Matches 980; Conservative 0; Mismatches 494; Indels 51; Gaps 3;

Qy 880 CGGGCACTTCTCTTCAAGCTGG---AGGCACCTCCCGCTCTTCTGTGACCAACATCACC 936
Db 317 CAGGCACTTCTATTCAACAGGACAGGACAGGACGCCACTGTTCAAGTACGCAACCCC 376

Qy 937 AGGTACCCACTGACGTCCAGACAGAGTACTCTCTCCGCCCGGACCCCTGCGCCGAG 996
Db 377 GGGATACACATGATGATCTGCTCTGTTTATCTCCGCTTACCGGCCACTTCTCTAGAAA 436

Qy 997 CACCTTCGCGCGCGCGCTTTAACTCAAGAGCCCTCAAGTACTGTAACCTGGAAGTG 1056
Db 437 CACCTATCAAGAGTCTTTTAAATCAAGAGTCTTCAAGATCTGAGCTGGAGGTG 496

Qy 1057 CGCAGCCCTGAGGCCATCTCTATCAGCCACTCTGTGTCATCTCTCTGTCATCTTT 1116
Db 497 CACCGCACTGTGTCTGTAGGGTCTCAGTGTCTCTGGCCATCTCTCTCTCTATTTAT 556

Qy 1117 GGCATGCACTGTTGGCTTAACCTGACCTGACCGCTGACCGATGAGGGGAGATGATGA 1176
Db 557 AGCAATGATCTATTTGGCTCAACTGGCACTTACAGACAGCGAATAAGACATCGA 616

Qy 1177 GATCAGGAGACACACCCAGCAGTGGCCCTGTGCCAACCGAGCTCTCCCTATACCCCTC 1236
Db 617 GA-----ATGGAAGTGAATTCTGACACCGTGCCCAACACACTGTATCGTTACTTC 670

Qy 1237 AGGGGCACTGGCTTAGAGACCCCTGACAGAAAGGCAAGAACACAGAGGAAGCC 1296
Db 671 TGGCGACA-----ATGCAAAAT 688

Qy 1297 CAGTAGTTTCTTCCAGAGACAGTTCATAGATTCTGGAGAAATGATGTGGGAAGCG 1356
Db 689 AGGTGGATTTACATAGAAATACACCATAGATTCCGGAGAACTTGATATTGGCCGGAG 748

Qy 1357 AGCTCCCAAGATTCCTCTCTGCACTTTCTGGAGATCTCAAGATGTTTATAGACCATCC 1416
Db 749 AGCAATTCAGAGAGTTCCCGCCGGATCTTCTGAGATCGAGCTCTTTATTGATCAGCC 808

Qy 1417 TGTCATCTGAATCAATGTCTCTGGGAAAGGAGCCCTGTTGGCAATTTATGGCAG 1476
Db 809 ACAGTTTCTTAAGTTCAACATCTCTCTTCAAGAGATGCAATTCATCGAGTGTACGCCG 868

Qy 1477 AAAAGGCTCCCTCTTACATACACAGTTTGAATTTGTGAGCTCTCTGATGGCAGGAG 1536
Db 869 GAAGGGCTTACCGGCTTCCCATACTAGTACGACTTTGTGAACTACTGGATGTTAGCAG 928

Qy 1537 GCTCTAACCCAGAGCGGAGCTTAGAGGGAGCCCGCCGACGTCTCGGGAACTGT 1596
Db 929 GTTAATTCAGAGAGCAGCGGAACCTGGTGGAGTCCGAAAGAGCCGGCGGAGCGAG 988
```

```
Qy 1597 GCCCCCTCCAGCCATGAGACAGAGCTTCAATCCAGTATTTGGATTCAGAAATCTGGCACTT 1656
Db 989 ATCTGTAGCTGACGAGAGCTGGCTTCAATCCAGTACTTGGATTCAGAAATCTGGCACTT 1648

Qy 1657 GCGTTTTTACAAATGACGGAAGAGTCAAGAGTGGTTTCTTTCTACCACTGCAATGA 1716
Db 1049 GCGTTTTTATAACGACGGGAAAAACCCAGAGCAGTCTCCTTTAAACGATCGTTATAGA 1108

Qy 1717 GTCGTGATTAACCTGCCCCAGCACTGCTATGGAATGTTGACTGATCTCTGGACCTG 1776
Db 1109 GTCGTGATGGAATGCCCCCGAAATGCCATGGAATGGAAGTGTGTTTCTGGAATG 1168

Qy 1777 CCACTGCTTCTCGGTTTCTTGGGCCCCGACGTGTGGCAGAGCTCTCTGCCCCCTGCTG 1836
Db 1169 CCAATGTTTCCCGGTTTCTAGGTCGGATGTTTCAAGAGCAGCTGTCCGGTGTCTG 1228

Qy 1837 TAGCGGAATGCGCAATACATGAAGGAGATGCTTGTGCCACAGTGGCTGGAAGCGC 1896
Db 1229 TAGTGGCAACGGGCAATACTCCAAGGGCCGCTGCTCTGCTTCAAGTGGTGGAGGGCAC 1288

Qy 1897 TGAGTGGATGTGCCCAACCAACAGTGTATCGATGTGGCCTGCGAGCAACCATGGCACTG 1956
Db 1289 CGAGTGTGACGTGCGAGCAGCCAGTGCATTTGACCCGAGTGGGGGTCTGGGATTTG 1348

Qy 1957 CATCAGGGGACCTGCACTCTGCAACCTTGGCTACAAGGGCAGAGCTGTGAGGAAGTGA 2016
Db 1349 CATCATGGGCTTTTGGCTTGTAACTCGGGATACAAAGAGAAACCTGTGAGGAAGCGGA 1408

Qy 2017 CTGATGAGACCCACATGTTCAAGCCGGGTGTCTGCTGAGAGCGCAATGCCATTCCTT 2076
Db 1409 CTGCTAGACCTGGATGTTCTAATACGGGGTGTGATCCATGGGAATGTCACTGCAA 1468

Qy 2077 TGTGGGATGGGGAGGACCAACTGCGAGACCCCGAGGCGCCACATGTTTAGACCAATGTTT 2136
Db 1469 TCCAGGCTGGGGTGGCAGCAACTGTGAATACTTGAAGACTATGTGTCAGACCACTGCTC 1528

Qy 2137 AGGCCAGGAACCTTCTCCCGGACACCGGGCTTTGCACTGTGACCCAGCTGGAGTGG 2196
Db 1529 AGGCCACGGGACTTACCTTCAAGAAAGCGGCTCCTGCACTTGGACACCAATTTGGACTG 1588

Qy 2197 ACAGACTGTTCTATCGAGATCTGCTGCGCACTGTGTGGCCATGGCGTGTGCGTAGG 2256
Db 1589 CCCGACTGCTCAATGAATATGTTCACTGGAATGCGGGCTCACACGGCTCTGCAATGG 1648

Qy 2257 GGGCACTGCGCTGCGAGGATGCTGGATGGGGGAGCCCTGCGACAGCGGGCTGCCA 2316
Db 1649 GGGCTCTGCTGCTGTGAAGAGGCTGGACCGGCGCGCGTGTAAATCAGAGAGCTTGCCA 1708

Qy 2317 CCCGCTGTGCGGAGCATGGGACCTGCCGACGCAAGTGGAGTGGAGCCCTGGCTG 2376
Db 1709 CCTGCTGTGCTGAGACAGGGAGCTGCAAGGACGGCAAGTGGAGTGGAGCCAGATG 1768

Qy 2377 GAATGGGGAACACTGCACCAATCGCT 2401
Db 1769 GAAAGGAGCACTGCACAAATGGT 1793
```

Search completed: June 25, 2004, 02:14:52
Job time : 1822 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 24, 2004, 17:03:54 ; Search time 2140 Seconds
(without alignments)
17883.374 Million cell updates/sec

Title: US-10-029-020-13
Perfect score: 8354
Sequence: 1 gtttggtggtgagagac.....ccggagggtgacagagaggac 8354

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3017426 seqs, 229054650 residues
Total number of hits satisfying chosen parameters: 6034852

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8354	100.0	8354	13	US-10-383-201-43
2	8354	100.0	8354	13	US-10-029-020-13
3	7991.2	95.7	8355	13	US-10-383-201-55
4	7799.2	93.1	8438	13	US-10-042-865-1
5	3502.2	41.9	8645	17	US-10-038-854-37
6	3473.8	41.6	8675	17	US-10-038-854-35
7	3422	41.0	8473	17	US-10-038-854-39
8	3398.8	40.7	8487	17	US-10-038-854-41
9	3397.8	40.7	8689	9	US-09-808-602-78
10	3397.8	40.7	8689	10	US-09-800-198-66
11	3385	40.5	9826	9	US-09-808-602-7
12	3385	40.5	9826	10	US-09-800-198-7
13	3384.8	40.5	8797	9	US-09-808-602-74
14	3384.8	40.5	8797	9	US-09-808-602-77

15	3384.8	40.5	8797	10	US-09-800-198-62
16	3384.8	40.5	8797	10	US-09-800-198-65
17	3351.4	40.1	9729	9	US-09-808-602-12
18	3351.4	40.1	9729	10	US-09-800-198-12
19	3258	39.0	12880	16	US-10-295-027-927
20	3165	37.9	8695	16	US-10-144-194A-81
21	3158.4	37.8	8575	13	US-10-072-012-143
22	3070.8	36.8	8409	9	US-09-808-602-79
23	3070.8	36.8	8409	10	US-09-800-198-67
24	3068.6	36.7	9058	16	US-10-144-194A-79
25	2533	30.3	6560	9	US-09-808-602-76
26	2533	30.3	6560	10	US-09-800-198-64
27	1708.8	20.5	3111	9	US-09-773-517-12
28	1708.8	20.5	3111	9	US-09-792-025-12
29	1708.8	20.5	3111	9	US-09-849-868-12
30	1708.8	20.5	3111	15	US-10-290-578-1
31	1708.8	20.5	3111	15	US-10-453-183-12
32	1673.6	20.0	1680	15	US-10-290-578-3
33	1560.4	18.7	3614	13	US-10-342-887-1743
34	1560.4	18.7	3614	13	US-10-172-118-1743
35	1378.2	16.5	1399	13	US-10-383-201-49
36	1378.2	16.5	1399	13	US-10-383-201-59
37	1377.2	16.5	1476	13	US-10-383-201-41
38	1375.2	16.5	1392	13	US-10-383-201-45
39	1375.2	16.5	1392	13	US-10-383-201-53
40	1309.4	15.7	1371	13	US-10-383-201-51
41	1185.6	14.2	2387	15	US-10-290-578-11
42	833	10.0	3217	15	US-10-198-846-13976
43	807.2	9.7	829	13	US-10-383-201-47
44	807.2	9.7	829	13	US-10-383-201-61
45	744	8.9	802	13	US-10-383-201-57

ALIGNMENTS

RESULT 1

US-10-383-201-43
; Sequence 43, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: Curasequest version 0.1
; SEQ ID NO 43
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

Query Match									
NAME/KEY: CDS									
LOCATION: (35)...(8341)									
US-10-383-201-43									
Best Local Similarity 100.0%; Score 8354; DB 13; Length 8354;									
Matches 8354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	GT	TT	GT	GT	GT	GT	GT	GT
DB	1	GT	TT	GT	GT	GT	GT	GT	GT
QY	61	CC	CT	CT	CT	CT	CT	CT	CT
DB	61	CC	CT	CT	CT	CT	CT	CT	CT
QY	121	CG	AG	GG	GG	GG	GG	GG	GG
DB	121	CG	AG	GG	GG	GG	GG	GG	GG
QY	181	CC	AG	GG	GG	GG	GG	GG	GG
DB	181	CC	AG	GG	GG	GG	GG	GG	GG
QY	241	GG	AA	TT	CT	CT	CT	CT	CT
DB	241	GG	AA	TT	CT	CT	CT	CT	CT
QY	301	GC	CC	CT	CT	CT	CT	CT	CT
DB	301	GC	CC	CT	CT	CT	CT	CT	CT
QY	361	GG	GG	CT	CT	CT	CT	CT	CT
DB	361	GG	GG	CT	CT	CT	CT	CT	CT
QY	421	GC	CT	CT	CT	CT	CT	CT	CT
DB	421	GC	CT	CT	CT	CT	CT	CT	CT
QY	481	TT	CA	AT	CT	CT	CT	CT	CT
DB	481	TT	CA	AT	CT	CT	CT	CT	CT
QY	541	CT	GC	AG	AA	CC	CG	CG	CG
DB	541	CT	GC	AG	AA	CC	CG	CG	CG
QY	601	CA	AC	AG	CA	CC	CG	CG	CG
DB	601	CA	AC	AG	CA	CC	CG	CG	CG
QY	661	CA	AC	CG	CG	CG	CG	CG	CG
DB	661	CA	AC	CG	CG	CG	CG	CG	CG
QY	721	CC	AG	AG	CC	CG	CG	CG	CG
DB	721	CC	AG	AG	CC	CG	CG	CG	CG
QY	781	CAG	AA	CT	AG	CA	TT	AG	CA
DB	781	CAG	AA	CT	AG	CA	TT	AG	CA
QY	841	CA	TT	CT	CG	CG	CG	CG	CG
DB	841	CA	TT	CT	CG	CG	CG	CG	CG
QY	901	TG	AG	GG	CA	CT	CT	CT	CT
DB	901	TG	AG	GG	CA	CT	CT	CT	CT
QY	961	AG	GT	AG	TT	CT	CT	CT	CT
DB	961	AG	GT	AG	TT	CT	CT	CT	CT

Db	961	AG	GT	AG	TT	CT	CT	CT	CT
QY	1021	CCT	CA	AG	AG	CC	CT	CT	CT
Db	1021	CCT	CA	AG	AG	CC	CT	CT	CT
QY	1081	CT	AG	CC	CT	CT	CT	CT	CT
Db	1081	CT	AG	CC	CT	CT	CT	CT	CT
QY	1141	CT	GC	AC	CT	CT	CT	CT	CT
Db	1141	CT	GC	AC	CT	CT	CT	CT	CT
QY	1201	TT	GC	CT	CT	CT	CT	CT	CT
Db	1201	TT	GC	CT	CT	CT	CT	CT	CT
QY	1261	TG	CA	GA	AG	CC	CT	CT	CT
Db	1261	TG	CA	GA	AG	CC	CT	CT	CT
QY	1321	TT	TC	AT	AG	AT	CT	CT	CT
Db	1321	TT	TC	AT	AG	AT	CT	CT	CT
QY	1381	CAC	TT	CT	GA	AT	CT	CT	CT
Db	1381	CAC	TT	CT	GA	AT	CT	CT	CT
QY	1441	TCT	CG	AA	AG	CC	CT	CT	CT
Db	1441	TCT	CG	AA	AG	CC	CT	CT	CT
QY	1501	AC	AG	TT	GA	AT	CT	CT	CT
Db	1501	AC	AG	TT	GA	AT	CT	CT	CT
QY	1561	CCT	AG	GG	AG	CC	CT	CT	CT
Db	1561	CCT	AG	GG	AG	CC	CT	CT	CT
QY	1621	CT	TC	AT	CC	AG	AT	CT	CT
Db	1621	CT	TC	AT	CC	AG	AT	CT	CT
QY	1681	GT	CA	AG	AG	CC	CT	CT	CT
Db	1681	GT	CA	AG	AG	CC	CT	CT	CT
QY	1741	CT	GT	AT	GG	CA	AT	CT	CT
Db	1741	CT	GT	AT	GG	CA	AT	CT	CT
QY	1801	CCC	CG	AC	TT	GG	CA	AT	CT
Db	1801	CCC	CG	AC	TT	GG	CA	AT	CT
QY	1861	AGG	CAG	AT	GG	CA	AT	CT	CT
Db	1861	AGG	CAG	AT	GG	CA	AT	CT	CT
QY	1921	GT	GT	AT	CG	AT	GG	CA	AT
Db	1921	GT	GT	AT	CG	AT	GG	CA	AT
QY	1981	CCC	TGG	CT	TA	AG	GG	CA	AT
Db	1981	CCC	TGG	CT	TA	AG	GG	CA	AT
QY	2041	CCG	GG	GT	CT	CT	CT	CT	CT
Db	2041	CCG	GG	GT	CT	CT	CT	CT	CT

QY 4261 CACAGATTAGCCATCAACCCAAATGGAACTCACTTTATGTCTCTCGACAAACAATGGT 4320
Db 4261 CACAGATTAGCCATCAACCCAAATGGAACTCACTTTATGTCTCTCGACAAACAATGGT 4320
QY 4321 CCTCAAAATCTCTGAAACCAACAGAGTGGCGATTTGTCGGCGAGGCGCCATGCACTGCCA 4380
Db 4321 CCTCAAAATCTCTGAAACCAACAGAGTGGCGATTTGTCGGCGAGGCGCCATGCACTGCCA 4380
QY 4381 GGTCCCTGGATGACACATCTCTGTAAGAGAGTGGCCATCCAGCAACCCCTGAGTC 4440
Db 4381 GGTCCCTGGATGACACATCTCTGTAAGAGAGTGGCCATCCAGCAACCCCTGAGTC 4440
QY 4441 AGCCACCGCTTTGGCTGTTTCAACAATGGGTCCTGTATATTGCTGAGACTGATGAGAA 4500
Db 4441 AGCCACCGCTTTGGCTGTTTCAACAATGGGTCCTGTATATTGCTGAGACTGATGAGAA 4500
QY 4501 AAAGATCAACCGCATAGGAGGTACCACTAGTGGAGATCTACTGTTGCTGGGC 4560
Db 4501 AAAGATCAACCGCATAGGAGGTACCACTAGTGGAGATCTACTGTTGCTGGGC 4560
QY 4561 CCCAGTGGCTGACACTGTAATAATGATGCAACTGTGATTGTTTCTGGAGACCATGG 4620
Db 4561 CCCAGTGGCTGACACTGTAATAATGATGCAACTGTGATTGTTTCTGGAGACCATGG 4620
QY 4621 TTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGTGTGCTGATGGGA 4680
Db 4621 TTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGTGTGCTGATGGGA 4680
QY 4681 GCTTACGTGGCCGACCTTGGACATCCGAAATCGGTTTATCCGGAAGACAGCTTT 4740
Db 4681 GCTTACGTGGCCGACCTTGGACATCCGAAATCGGTTTATCCGGAAGACAGCTTT 4740
QY 4741 CCTCAACACCCAGAACATGATGAGCTGTCTTCAACAAATGACAGAGCTCTATCTGTT 4800
Db 4741 CCTCAACACCCAGAACATGATGAGCTGTCTTCAACAAATGACAGAGCTCTATCTGTT 4800
QY 4801 TGATACACCGGAGCACTGTACCCAAAGCTGCCACAGGAGACTACCTGTACAA 4860
Db 4801 TGATACACCGGAGCACTGTACCCAAAGCTGCCACAGGAGACTACCTGTACAA 4860
QY 4861 CTTCACTTACACTGGGAGCGGACATCACACTATCACAGACAAATGGCAACATGTT 4920
Db 4861 CTTCACTTACACTGGGAGCGGACATCACACTATCACAGACAAATGGCAACATGTT 4920
QY 4921 AAATGTCGCGAGACTTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGT 4980
Db 4921 AAATGTCGCGAGACTTACTGGGATGCCCTCTGGCTGGTGGTCCAGATGGCCAGT 4980
QY 4981 GTACTGGTGACCATGGGACCAACAGTGCATCAAGAGTGTACACACAGGACCA 5040
Db 4981 GTACTGGTGACCATGGGACCAACAGTGCATCAAGAGTGTACACACAGGACCA 5040
QY 5041 GTTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAGCAATGAAA 5100
Db 5041 GTTGGCCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAGCAATGAAA 5100
QY 5101 CGGATGACACATTTATGAGTACGACAGCTTGGCGGCTGACAAATGTGACCTTCCC 5160
Db 5101 CGGATGACACATTTATGAGTACGACAGCTTGGCGGCTGACAAATGTGACCTTCCC 5160
QY 5161 TACTGGCCAGTGACAGTTCGGAAGTATACAGACAGTTCACTGATGTCAGGTAGA 5220
Db 5161 TACTGGCCAGTGACAGTTCGGAAGTATACAGACAGTTCACTGATGTCAGGTAGA 5220
QY 5221 GACCTCAGCAAGATGATGTACCATACCAACCTGTCTGCTCAGGCGCCTTCTA 5280
Db 5221 GACCTCAGCAAGATGATGTACCATACCAACCTGTCTGCTCAGGCGCCTTCTA 5280
QY 5281 CACACTGTCGAACCAAGTCGGAACAGCTACTACTCATCTCGGGGCGATGGCTTCTTGG 5340
Db 5281 CACACTGTCGAACCAAGTCGGAACAGCTACTACTCATCTCGGGGCGATGGCTTCTTGG 5340
QY 5341 GCTGCTGCTGCCAAACCGCATGGAGTGGCGCTGCAGACTGAGCCCACTTGTGCTGG 5400

Db 5341 GCTGCTGCTGCCAAACCGCATGGAGTGGCGCTGCAGACTGAGCCCACTTGTGCTGGTGG 5400
QY 5401 CACCGTCAACCCCAACCGTGGGCAAGAGAAATGTCACTGCGCCCATCGACAACGGCCTCAA 5460
Db 5401 CACCGTCAACCCCAACCGTGGGCAAGAGAAATGTCACTGCGCCCATCGACAACGGCCTCAA 5460
QY 5461 CCTGCTGGAGTGGCGCCAGCCAAAGAGAGGCTCGGGGCGAGTCACTGCTTTTGGCG 5520
Db 5461 CCTGCTGGAGTGGCGCCAGCCAAAGAGAGGCTCGGGGCGAGTCACTGCTTTTGGCG 5520
QY 5521 CCGGCTGGGGTGCACAAACCGAAATCTCTTCTGGAATTTGATCGCGTAAACAGCAC 5580
Db 5521 CCGGCTGGGGTGCACAAACCGAAATCTCTTCTGGAATTTGATCGCGTAAACAGCAC 5580
QY 5581 AGAGAAGATCTATGATGACCAACCGAAATTCACCCCTTCGGATTTCTGTACGACGCGGG 5640
Db 5581 AGAGAAGATCTATGATGACCAACCGAAATTCACCCCTTCGGATTTCTGTACGACGCGGG 5640
QY 5641 CGGGCCAGCCTCTGGTCAACCCAGCAGGCTGAATGGTGTCAACCGTGAATCTCTCCC 5700
Db 5641 CGGGCCAGCCTCTGGTCAACCCAGCAGGCTGAATGGTGTCAACCGTGAATCTCTCCC 5700
QY 5701 TGGGGTTCATTTGCTGGCATCCAGAGGCGATCATGTCTGAAAGAAATGGAATACGACCA 5760
Db 5701 TGGGGTTCATTTGCTGGCATCCAGAGGCGATCATGTCTGAAAGAAATGGAATACGACCA 5760
QY 5761 GCGGGCGCATCATCTCAGGATCTTCTGCTGATGGGAAGACATGAGCTACACATACTT 5820
Db 5761 GCGGGCGCATCATCTCAGGATCTTCTGCTGATGGGAAGACATGAGCTACACATACTT 5820
QY 5821 AGAGAAGTCCATGGTGTCTTACTACAGCAGGAGCAGTATATCTTTGAGTTCGACAA 5880
Db 5821 AGAGAAGTCCATGGTGTCTTACTACAGCAGGAGCAGTATATCTTTGAGTTCGACAA 5880
QY 5881 GAATGACCGCTCTTCTGTGACGATGCCCAAGTGGCGGCGGACAGACTAGAGCAT 5940
Db 5881 GAATGACCGCTCTTCTGTGACGATGCCCAAGTGGCGGCGGACAGACTAGAGCAT 5940
QY 5941 CCGCTCAGTGGGCTACTACAGAAATCTATCAGCCCTCGAGGGCAATGCCTCAGTCAT 6000
Db 5941 CCGCTCAGTGGGCTACTACAGAAATCTATCAGCCCTCGAGGGCAATGCCTCAGTCAT 6000
QY 6001 ACAGACTTCACTGAGGATGGCACTCTTACACCTTCTACCTGGGCACTGGCCGAG 6060
Db 6001 ACAGACTTCACTGAGGATGGCACTCTTACACCTTCTACCTGGGCACTGGCCGAG 6060
QY 6061 GGTGATATACAAATGATGGCAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAACAA 6120
Db 6061 GGTGATATACAAATGATGGCAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCAACAA 6120
QY 6121 GGTGATTTTCACTTACGAGAGCGGAGGATGCTGAAAGCAGTCAACCTACAGAAATGA 6180
Db 6121 GGTGATTTTCACTTACGAGAGCGGAGGATGCTGAAAGCAGTCAACCTACAGAAATGA 6180
QY 6181 GGGCTTCACTGACCATCCGCTACCGTACAGATGGGCGGCTGATTCAGCGACAGATCTT 6240
Db 6181 GGGCTTCACTGACCATCCGCTACCGTACAGATGGGCGGCTGATTCAGCGACAGATCTT 6240
QY 6241 CCGCTTCACTGAGGAAGGCTGTCACCGCCGCTTTGACTTACAACTATGACAAACAGCTT 6300
Db 6241 CCGCTTCACTGAGGAAGGCTGTCACCGCCGCTTTGACTTACAACTATGACAAACAGCTT 6300
QY 6301 CCGGCTGACAGCATGAGGCTGTGATCAAGAGACCCCACTGCCATTCATCTATCG 6360
Db 6301 CCGGCTGACAGCATGAGGCTGTGATCAAGAGACCCCACTGCCATTCATCTATCG 6360
QY 6361 CTATGATGATGTGTGAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCTTACTATGA 6420
Db 6361 CTATGATGATGTGTGAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCTTACTATGA 6420
QY 6421 CATTAACAGATCATCAACAGCTGTCTGACCCACCAAGCAATTTTGTGATGATATGG 6480

Db 6421 CATTAACCATGATCATACACAGAGTGTCTATGACCCACCAAGCAATTTTGATGCATATGG 6480
 Qy 6481 CAGATGAAGAAAGTGCAGTATGAGATCTTCCCTCGCTCATCTACTGTGATGACCTCCA 6540
 Db 6481 CAGATGAAGAAAGTGCAGTATGAGATCTTCCCTCGCTCATCTACTGTGATGACCTCCA 6540
 Qy 6541 GTATGATAACATGGGGGAGTGTGAGAGAGAGCTGAGGTAGGACCTTACGCCAATAC 6600
 Db 6541 GTATGATAACATGGGGGAGTGTGAGAGAGAGCTGAGGTAGGACCTTACGCCAATAC 6600
 Qy 6601 CACTCGCTACTCTATGATGATGCTGACGGCCAGCTTGCAGACAGTCTTCCATCAATGA 6660
 Db 6601 CACTCGCTACTCTATGATGATGCTGACGGCCAGCTTGCAGACAGTCTTCCATCAATGA 6660
 Qy 6661 CAAGCCACTCTGGGCTTACAGCTACGACCTCAATGGAACTTCACTTACTGAGCCCTGG 6720
 Db 6661 CAAGCCACTCTGGGCTTACAGCTACGACCTCAATGGAACTTCACTTACTGAGCCCTGG 6720
 Qy 6721 GAACAGTGCAGGCTCACACCACTACCGGTATGATGATCCGCGACCGCATCTCGGCTGG 6780
 Db 6721 GAACAGTGCAGGCTCACACCACTACCGGTATGATGATCCGCGACCGCATCTCGGCTGG 6780
 Qy 6781 TGACGTGCATACAAAGTGGATGAGGATGGCTTCTTGAGGACGGGGGGGTGATATCTT 6840
 Db 6781 TGACGTGCATACAAAGTGGATGAGGATGGCTTCTTGAGGACGGGGGGGTGATATCTT 6840
 Qy 6841 TGAGTCAAACTCAGCTGGCTGCTCAATCAAGGCTTCAACCGGCTGGAGCTGGAGTGT 6900
 Db 6841 TGAGTCAAACTCAGCTGGCTGCTCAATCAAGGCTTCAACCGGCTGGAGCTGGAGTGT 6900
 Qy 6901 CAGGTACCGCTACGATGGCTGGGGGGCGCGTGTCCAGCAAGAGCAGCCACAGCCACCA 6960
 Db 6901 CAGGTACCGCTACGATGGCTGGGGGGCGCGTGTCCAGCAAGAGCAGCCACAGCCACCA 6960
 Qy 6961 CCTGCAGTCTTCTATGACAGCTGACCAACCCCAACCAAGGTACCCACCTGTGACACCA 7020
 Db 6961 CCTGCAGTCTTCTATGACAGCTGACCAACCCCAACCAAGGTACCCACCTGTGACACCA 7020
 Qy 7021 CTCAGCTCTGAGATCACTTCTCTACTACGACTTGCAGGACACCTCTTTCGCAATGA 7080
 Db 7021 CTCAGCTCTGAGATCACTTCTCTACTACGACTTGCAGGACACCTCTTTCGCAATGA 7080
 Qy 7081 GCTGAGCAGTGTGATGATGATTTTACATAGCTTGTGACAAATCCGAGCCCTCTTCTGT 7140
 Db 7081 GCTGAGCAGTGTGATGATGATTTTACATAGCTTGTGACAAATCCGAGCCCTCTTCTGT 7140
 Qy 7141 CTTTGTGGAACAGCTTGTGATGATCAAGCAAACTCTGTACACAGCTATGGGAGATCTA 7200
 Db 7141 CTTTGTGGAACAGCTTGTGATGATCAAGCAAACTCTGTACACAGCTATGGGAGATCTA 7200
 Qy 7201 CATGGATACCAACCCCACTTTCAGATCATATAGGCTACCATGGTGGCTCTTATGATCC 7260
 Db 7201 CATGGATACCAACCCCACTTTCAGATCATATAGGCTACCATGGTGGCTCTTATGATCC 7260
 Qy 7261 ACTCACAAGCTTGTCCACATGGGCGGCGAGATTTATGATGCTGGCGGAGCGCTGGAC 7320
 Db 7261 ACTCACAAGCTTGTCCACATGGGCGGCGAGATTTATGATGCTGGCGGAGCGCTGGAC 7320
 Qy 7321 TAGCCACAGCAGCTGTGGAAGCAGCTTATGATGACCAAGCTCATGCTTTTAACTCT 7380
 Db 7321 TAGCCACAGCAGCTGTGGAAGCAGCTTATGATGACCAAGCTCATGCTTTTAACTCT 7380
 Qy 7381 CTATATGTTCAAAAACAAACCCCATCAGCAACTCCAGGACATCAAGTGTCTTATGAC 7440
 Db 7381 CTATATGTTCAAAAACAAACCCCATCAGCAACTCCAGGACATCAAGTGTCTTATGAC 7440
 Qy 7441 AGATGTTAAACAGCTGGCTGTCACTTTGATTTCCAGTACCAAAAGTATCCCTGGTTA 7500
 Db 7441 AGATGTTAAACAGCTGGCTGTCACTTTGATTTCCAGTACCAAAAGTATCCCTGGTTA 7500
 Qy 7501 TCCAAAACAGACATGATGCCATGGAAACCTCTCTAGAGCTCATCCACACAGATGAA 7560
 Db 7501 TCCAAAACAGACATGATGCCATGGAAACCTCTCTAGAGCTCATCCACACAGATGAA 7560

RESULT 2

US-10-029-020-13
 ; Sequence 13, Application US/10029020
 ; Publication No. US2004003971A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gangoli et al.
 ; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-225
 ; CURRENT APPLICATION NUMBER: US/10/029,020
 ; CURRENT FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/256,704
 ; PRIOR FILING DATE: 2000-12-19
 ; PRIOR APPLICATION NUMBER: 60/311,590
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: 60/257,314

Qy 7561 AACGACAGGTTGGGACAAACAGCAAGTCTATCTCTGGGGTACAGTGTGAAGTACAGAAGCA 7620
 Db 7561 AACGACAGGTTGGGACAAACAGCAAGTCTATCTCTGGGGTACAGTGTGAAGTACAGAAGCA 7620
 Qy 7621 GCTCAAGGCCCTTTGTACCTTTAGAACGGTTTGACAGCTCTATGCTCCCAATCACCAG 7680
 Db 7621 GCTCAAGGCCCTTTGTACCTTTAGAACGGTTTGACAGCTCTATGCTCCCAATCACCAG 7680
 Qy 7681 CTGCCAGCAGGCTCCAAAGACCAAGAGTTTGCATTCGAGCGGCTCAGTCTTTGGCAAGGG 7740
 Db 7681 CTGCCAGCAGGCTCCAAAGACCAAGAGTTTGCATTCGAGCGGCTCAGTCTTTGGCAAGGG 7740
 Qy 7741 GGTCAAGTTTGGCTTGAAGGATGGCCGAGTGACACAGACATCATCAGTGTGGCCCAATGA 7800
 Db 7741 GGTCAAGTTTGGCTTGAAGGATGGCCGAGTGACACAGACATCATCAGTGTGGCCCAATGA 7800
 Qy 7801 GGTGGGGAAGGGTGTGCTGCATCTTGAACCATCTTGAACCATCTTGAACCATCTTGAACCATCTT 7860
 Db 7801 GGTGGGGAAGGGTGTGCTGCATCTTGAACCATCTTGAACCATCTTGAACCATCTTGAACCATCTT 7860
 Qy 7861 CACCATTTGATGGGTGGATACCCATTACTTTTGTGAACCCAGACCTTCCAGAAAGTGCCT 7920
 Db 7861 CACCATTTGATGGGTGGATACCCATTACTTTTGTGAACCCAGACCTTCCAGAAAGTGCCT 7920
 Qy 7921 GGCCATCTGGGCTCAGTGGGGGGCGGCAACCTTGAGAAATGGGGTCAACGTCACCTGT 7980
 Db 7921 GGCCATCTGGGCTCAGTGGGGGGCGGCAACCTTGAGAAATGGGGTCAACGTCACCTGT 7980
 Qy 7981 GTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA 8040
 Db 7981 GTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA 8040
 Qy 8041 GTACGGGCACTGTCTTGAACACAGCTACGGGACACCTTGGATGAGGAGAGGCGACG 8100
 Db 8041 GTACGGGCACTGTCTTGAACACAGCTACGGGACACCTTGGATGAGGAGAGGCGACG 8100
 Qy 8101 GGTCTGAGCTGGCCCGGCGGAGAGGCGTGGCCAAAGCTGGCCGCGGAGGAGGCGAG 8160
 Db 8101 GGTCTGAGCTGGCCCGGCGGAGAGGCGTGGCCAAAGCTGGCCGCGGAGGAGGCGAG 8160
 Qy 8161 ACTCGGGAAGGGGAGGAGGCGCTGGGCGCTGGACAGAGGCGGAGAGGAGGAGGAGGAGG 8220
 Db 8161 ACTCGGGAAGGGGAGGAGGCGCTGGGCGCTGGACAGAGGCGGAGAGGAGGAGGAGGAGG 8220
 Qy 8221 GAGCACAGGGGGGTGCAAGCTACGAGGCTTTTCTGATCTCTGTCGAGCAGTACCC 8280
 Db 8221 GAGCACAGGGGGGTGCAAGCTACGAGGCTTTTCTGATCTCTGTCGAGCAGTACCC 8280
 Qy 8281 AGAATGTGACAGGCGCAACATCCATCTTATGAGACAGAGCGGAGATGGGCCGAG 8340
 Db 8281 AGAATGTGACAGGCGCAACATCCACTTCTATGAGACAGAGCGGAGATGGGCCGAG 8340
 Qy 8341 GTGACAGAGAGGAC 8354
 Db 8341 GTGACAGAGAGGAC 8354

```

; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 8354
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-029-020-13

Query Match      100.0%; Score 8354; DB 13; Length 8354;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 8354; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTTTGTGGATGTGAGAGCGCGGCGGAGGCGCATGGAGTGAAGAGAGAGAGCGCTTA 60
Db 1 GTTTGTGGATGTGAGAGCGCGGCGGAGGCGCATGGAGTGAAGAGAGAGAGCGCTTA 60

Qy 61 CCGCTCCCTGACCGCGCGCGCGAGCGCGCGTACACAGCTCGTCCGCGAGAG 120
Db 61 CCGCTCCCTGACCGCGCGCGCGAGCGCGCGTACACAGCTCGTCCGCGAGAG 120

Qy 121 CGAGAGGGCAAGCCCGCGAGAAATGTACAGCTCCAGCGAGACCTGAAGGCTTACGA 180
Db 121 CGAGAGGGCAAGCCCGCGAGAAATGTACAGCTCCAGCGAGACCTGAAGGCTTACGA 180

Qy 181 CCAGGACCGCGCTAGCTATGGCAGCGCGCTCAAGGACATGTGCGCGAGAGCGGA 240
Db 181 CCAGGACCGCGCTAGCTATGGCAGCGCGCTCAAGGACATGTGCGCGAGAGCGGA 240

Qy 241 GGAATTCGCGGACAGGTCGCAATTCACCTCGCGGAGCTGGGCTGGAAGAAGTAAC 300
Db 241 GGAATTCGCGGACAGGTCGCAATTCACCTCGCGGAGCTGGGCTGGAAGAAGTAAC 300

Qy 301 GCGCCCTCAGGAGCCCTGTACCGGACAGACATGGCGCTGCCCAATGGCGTACTCCAT 360
Db 301 GCGCCCTCAGGAGCCCTGTACCGGACAGACATGGCGCTGCCCAATGGCGTACTCCAT 360

Qy 361 GGGGGCTGGTCTGTATCCGACATGAGGCTGACCGGTCTGCTCCCTGAGCACCCTGT 420
Db 361 GGGGGCTGGTCTGTATCCGACATGAGGCTGACCGGTCTGCTCCCTGAGCACCCTGT 420

Qy 421 GCGTCTGTGGGCGGAGCAACGCTCAGGCGGAGCTGCGGTCTGCTGCTGCTCAGCGGCGCAA 480
Db 421 GCGTCTGTGGGCGGAGCAACGCTCAGGCGGAGCTGCGGTCTGCTGCTGCTCAGCGGCGCAA 480

Qy 481 TTCCAAATCTCAGCTACCGACCGAGGATGAACATCTGAGCTGATCATCCGGGCGG 540
Db 481 TTCCAAATCTCAGCTACCGACCGAGGATGAACATCTGAGCTGATCATCCGGGCGG 540

Qy 541 CTTGAGAAACAGCGCGGCTCCGAGCGCGCGCGCGCTCTGCGACGCCCAACACCC 600
Db 541 CTTGAGAAACAGCGCGGCTCCGAGCGCGCGCGCGCTCTGCGACGCCCAACACCC 600

Qy 601 CAACCCAGCAACCGCGCTCTCAATTAATCCCTGAAACCGGGGCACTTCAACCGGAGG 660
Db 601 CAACCCAGCAACCGCGCTCTCAATTAATCCCTGAAACCGGGGCACTTCAACCGGAGG 660

Qy 661 CAACCCAGCAACCGCGCTCTCAATTAATCCCTGAAACCGGGGCACTTCAACCGGAGG 720
Db 661 CAACCCAGCAACCGCGCTCTCAATTAATCCCTGAAACCGGGGCACTTCAACCGGAGG 720

```

```

721 CCAGGAGCCTCCACGCGCCAGGAGAACTGGCTCTCAACAGCAACATCCCTCTGGAGAC 780
721 CCAGGAGCCTCCACGCGCCAGGAGAACTGGCTCTCAACAGCAACATCCCTCTGGAGAC 780
781 CAGGAACCTAGGCAAGCAGCCATTTCTTAGGACATTCAGGACAACTCTTCTTAGATGGA 840
781 CAGGAACCTAGGCAAGCAGCCATTTCTTAGGACATTCAGGACAACTCTTCTTAGATGGA 840
841 CATTTCTGGCGCCTCCGCGCATGATGGGCTTACAGTGACGGGCACTTCTCTTCAAGCC 900
841 CATTTCTGGCGCCTCCGCGCATGATGGGCTTACAGTGACGGGCACTTCTCTTCAAGCC 900
901 TGGAGGACCTCCGCGCTCTTCTGCAACATCAACAGGTACCCATGACGTCCAGCAC 960
901 TGGAGGACCTCCGCGCTCTTCTGCAACATCAACAGGTACCCATGACGTCCAGCAC 960
961 AGTGTACTCTCCGCGCGCGAGCCCTGCGCGAGCACTTCCGCGCGCGCGCTTTAA 1020
961 AGTGTACTCTCTCCGCGCGCGAGCCCTGCGCGAGCACTTCCGCGCGCGCGCTTTAA 1020
1021 CCTCAAGAAGCCCTCCAAGTACTGTAAGTGAAGTGGCGAGCCCTGAGCGCCATCGTCAT 1080
1021 CCTCAAGAAGCCCTCCAAGTACTGTAAGTGAAGTGGCGAGCCCTGAGCGCCATCGTCAT 1080
1081 CTCAGCCACTCTGTGTCATCTCTGCTGGCATACTTTGCGCCATGCACCTGTTGGCCTAAA 1140
1081 CTCAGCCACTCTGTGTCATCTCTGCTGGCATACTTTGCGCCATGCACCTGTTGGCCTAAA 1140
1141 CTGCACTCTGAGCGGATGGAGGGGAGATGTATGAGATCAGCGAGGACACAGCCAGCAG 1200
1141 CTGCACTCTGAGCGGATGGAGGGGAGATGTATGAGATCAGCGAGGACACAGCCAGCAG 1200
1201 TTGCGCTGTGCGCAACGAGCTCTCCCTATACCTCTAGGGGGGAGCTGGCTTAGAGACCC 1260
1201 TTGCGCTGTGCGCAACGAGCTCTCCCTATACCTCTAGGGGGGAGCTGGCTTAGAGACCC 1260
1261 TGACAGAAAGGCAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1320
1261 TGACAGAAAGGCAAGGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1320
1321 TTTTCATAGATTTCTGGAGAAATTTGATGTGGAAAGCGAGCTCCAGAGAAATTCCTCTGG 1380
1321 TTTTCATAGATTTCTGGAGAAATTTGATGTGGAAAGCGAGCTCCAGAGAAATTCCTCTGG 1380
1381 CACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCACTGGAATTCATGTGTC 1440
1381 CACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCACTGGAATTCATGTGTC 1440
1441 TCTGGGAAAGGCAAGCCTCTGGTGGCAATTTATGGCAAAAAGGCGCTCCCTCTTCAATAC 1500
1441 TCTGGGAAAGGCAAGCCTCTGGTGGCAATTTATGGCAAAAAGGCGCTCCCTCTTCAATAC 1500
1501 ACAGTTTGAATTTCTGGAGCTCTGGATGGAGAGCTCTTAACCCAGAGAGCGCGGAG 1560
1501 ACAGTTTGAATTTCTGGAGCTCTGGATGGAGAGCTCTTAACCCAGAGAGCGCGGAG 1560
1561 CCTAGAGGGGACCCCGCGCGAGCTCTCGGGGAACTGTGCGCGCGCTCCAGCCATGAGACAG 1620
1561 CCTAGAGGGGACCCCGCGCGAGCTCTCGGGGAACTGTGCGCGCGCTCCAGCCATGAGACAG 1620
1621 CTTTCATCAGATTTTGGATTCAGGAATCTGGCACTTGGCTTTTTTAATGACGGAAGGA 1680
1621 CTTTCATCAGATTTTGGATTCAGGAATCTGGCACTTGGCTTTTTTAATGACGGAAGGA 1680
1681 GTCAAGAGTGTGTTTCTTTCTTCTCACCCTGCAATTTAGTTCGTTGATTAATGCGCCAGCAA 1740
1681 GTCAAGAGTGTGTTTCTTTCTTCTCACCCTGCAATTTAGTTCGTTGATTAATGCGCCAGCAA 1740
1741 CTGCTATGGCAATGGTGAAGTCACTCTCTGGGAGCTGCACTGCTTCTGCGGTTTCTGGG 1800
1741 CTGCTATGGCAATGGTGAAGTCACTCTCTGGGAGCTGCACTGCTTCTGCGGTTTCTGGG 1800
1801 CCCGAGCTGTGGCAGAGCCTCTGCGCGGCTGCTCTGTAGCGGAATGGCCCAATACATGAA 1860

```

Db 1801 CCCCAGCTGGCAGAGCCTCTGCCCCGCTGCTGTAGCGGAAATGGCCAAATACATGAA 1860
Qy 1861 AGCAGATGCTTGCCACAGTGGCTGGAAGGCGCTGAGTGCGCCACCAACCA 1920
Db 1861 AGCAGATGCTTGCCACAGTGGCTGGAAGGCGCTGAGTGCGCCACCAACCA 1920
Qy 1921 GTGTATCGATGTGGCCCTGCAGCAACATGCGACCTGCAATCAGCGGACCTGCAATCGCAA 1980
Db 1921 GTGTATCGATGTGGCCCTGCAGCAACATGCGACCTGCAATCAGCGGACCTGCAATCGCAA 1980
Qy 1981 CCTTGGCTCAAGGGCGAGAGCTGTGAGGAAGTGGACTGATGGAACCCACATGTTTCAGG 2040
Db 1981 CCTTGGCTCAAGGGCGAGAGCTGTGAGGAAGTGGACTGATGGAACCCACATGTTTCAGG 2040
Qy 2041 CCGGGGTGTCTGCGTGAGAGCGCAATGCCAATGCTTTGTGGGATGGGAGGACCAACTG 2100
Db 2041 CCGGGGTGTCTGCGTGAGAGCGCAATGCCAATGCTTTGTGGGATGGGAGGACCAACTG 2100
Qy 2101 CGAGACCCCGAGGCGCCATGCTTTAGACAGTGTTCAGGCCACGGAACTTCTCCCGGA 2160
Db 2101 CGAGACCCCGAGGCGCCATGCTTTAGACAGTGTTCAGGCCACGGAACTTCTCCCGGA 2160
Qy 2161 CACCGGCTTTGAGCTGTGACCCAGCTGGAAGTGGAGTGGAGTGTCTATCAGATCTG 2220
Db 2161 CACCGGCTTTGAGCTGTGACCCAGCTGGAAGTGGAGTGGAGTGTCTATCAGATCTG 2220
Qy 2221 TGTGCGGACTGTGGTGGCCATGCGCTGTGCGTAGGGGGCACTGCGCGCTGCGAGATGG 2280
Db 2221 TGTGCGGACTGTGGTGGCCATGCGCTGTGCGTAGGGGGCACTGCGCGCTGCGAGATGG 2280
Qy 2281 CTGGATGGGGGACCTGCGACCAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2340
Db 2281 CTGGATGGGGGACCTGCGACCAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2340
Qy 2341 CTGCGCGAGCGGCAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2400
Db 2341 CTGCGCGAGCGGCAAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2400
Qy 2401 TCATATCTGATAGGTTAGTTAAAGAGGTTGCTGAGTGGAGTGGAGTGGAGTGGAG 2460
Db 2401 TCATATCTGATAGGTTAGTTAAAGAGGTTGCTGAGTGGAGTGGAGTGGAGTGGAG 2460
Qy 2461 ATGTACCTTAGACCTGAACTGGTGGACCTGCTGCTGCGAGTGGAGTGGAGTGGAG 2520
Db 2461 ATGTACCTTAGACCTGAACTGGTGGACCTGCTGCTGCGAGTGGAGTGGAGTGGAG 2520
Qy 2521 CTGTGACACTTCCATGAGACTGCTGCGGTGACGAAAGACAAATGATGGAGTGGCCT 2580
Db 2521 CTGTGACACTTCCATGAGACTGCTGCGGTGACGAAAGACAAATGATGGAGTGGCCT 2580
Qy 2581 GGTGGACTGATGGAACCTGACTGCTGCTGCGAGTGGAGTGGAGTGGAGTGGAGTGG 2640
Db 2581 GGTGGACTGATGGAACCTGACTGCTGCTGCGAGTGGAGTGGAGTGGAGTGGAGTGG 2640
Qy 2641 CTTTGGCTCCCTTAACTCTGAGACTATCCAGGAGACAGGTCCCTGTGTCAAGCA 2700
Db 2641 CTTTGGCTCCCTTAACTCTGAGACTATCCAGGAGACAGGTCCCTGTGTGTCAAGCA 2700
Qy 2701 GAACCTACACTCTTCTATGACCGCATCAAGTTCCTGCGGGAGGAGACAGCAGCAT 2760
Db 2701 GAACCTACACTCTTCTATGACCGCATCAAGTTCCTGCGGGAGGAGACAGCAGCAT 2760
Qy 2761 AATCCCGGGAGAACCCCTTTGATGGAGGACATGCTTGTATTCGTCGCAAGTGT 2820
Db 2761 AATCCCGGGAGAACCCCTTTGATGGAGGACATGCTTGTATTCGTCGCAAGTGT 2820
Qy 2821 GACATCAGATGGAACCCCTTGTGTGAAATCAGTGTGTAATTAACCTCTCTT 2880
Db 2821 GACATCAGATGGAACCCCTTGTGTGAAATCAGTGTGTAATTAACCTCTCTT 2880
Qy 2881 TGGATATACATCAGAGGACAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 2940

Db 2881 TGGATATACATCAGAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGCATCTC 2940
Qy 2941 CATCATCTGCGGTTTCGAGCGGGACCTTTTCATACACAGAGACACACCTTGTGGCTGCC 3000
Db 2941 CATCATCTGCGGTTTCGAGCGGGACCTTTTCATACACAGAGACACACCTTGTGGCTGCC 3000
Qy 3001 ATGGGATCGCTTCTTTGTTCATGGAAACCATCATCATGAGACATGAGGAGAAATGAGATTCC 3060
Db 3001 ATGGGATCGCTTCTTTGTTCATGGAAACCATCATCATGAGACATGAGGAGAAATGAGATTCC 3060
Qy 3061 CAGCTGTGACCTGAGCAATTTTGGCCGCCCAACCCAGTGTCTCTCCATCCCACTGAC 3120
Db 3061 CAGCTGTGACCTGAGCAATTTTGGCCGCCCAACCCAGTGTCTCTCCATCCCACTGAC 3120
Qy 3121 GTCTTTCGCCAGCTCCTGTGCAGAGAAAGGCCCATTTGTGCCGGAATTCAGGCTTTTGA 3180
Db 3121 GTCTTTCGCCAGCTCCTGTGCAGAGAAAGGCCCATTTGTGCCGGAATTCAGGCTTTTGA 3180
Qy 3181 GGAGAAATCTCTATCTCTGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCGGACCCC 3240
Db 3181 GGAGAAATCTCTATCTCTGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCGGACCCC 3240
Qy 3241 TGGCTACAAATCTGTCTGAGGATCAGCTCACCACCCGACCATCCCTTCAACCTCAT 3300
Db 3241 TGGCTACAAATCTGTCTGAGGATCAGCTCACCACCCGACCATCCCTTCAACCTCAT 3300
Qy 3301 GAAGTGCACTCTGATGAGGAGGCGGCTCTTTCAGGAAAGTGGTTCGCTGACG 3360
Db 3301 GAAGTGCACTCTGATGAGGAGGCGGCTCTTTCAGGAAAGTGGTTCGCTGACG 3360
Qy 3361 CCCAGACCTGTCTATTTATTTTCATTTGGGACAGACAGCTCTACACCAAGAGTGT 3420
Db 3361 CCCAGACCTGTCTATTTATTTTCATTTGGGACAGACAGCTCTACACCAAGAGTGT 3420
Qy 3421 TGGGCTTTTCAGAAAGCTTTGTTTCCGTGGGTTATGAATATGAATCTCTCCAGATCTAAT 3480
Db 3421 TGGGCTTTTCAGAAAGCTTTGTTTCCGTGGGTTATGAATATGAATCTCTCCAGATCTAAT 3480
Qy 3481 CCTGTGGGAAAAGAAACAAACAGTCTCAGGGCTATGAATTCAGCGTCCAGCTTGG 3540
Db 3481 CCTGTGGGAAAAGAAACAAACAGTCTCAGGGCTATGAATTCAGCGTCCAGCTTGG 3540
Qy 3541 AGGATGGAGCCTTAGACAAACATCATGCGCTCAAAATTCAGGAGTATCTCTGCAAAAG 3600
Db 3541 AGGATGGAGCCTTAGACAAACATCATGCGCTCAAAATTCAGGAGTATCTCTGCAAAAG 3600
Qy 3601 GAATGGGAGAACCAAGTTTGTCTCAGCAGCTCCTGTCTATTTGGAGCATATGGCAA 3660
Db 3601 GAATGGGAGAACCAAGTTTGTCTCAGCAGCTCCTGTCTATTTGGAGCATATGGCAA 3660
Qy 3661 TGGGCGCGGAGAGCATCTCTGCGCCAGCTGCAACCGCTTGTCTGAGGCAACAGCT 3720
Db 3661 TGGGCGCGGAGAGCATCTCTGCGCCAGCTGCAACCGCTTGTCTGAGGCAACAGCT 3720
Qy 3721 CCTGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCTCTATGTGGGTGATTTCAA 3780
Db 3721 CCTGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCTCTATGTGGGTGATTTCAA 3780
Qy 3781 CTACATTAGAAGGATCTTCCCTCTGGAAATGTACCAACATCTTAGAGCTGAGGAAATA 3840
Db 3781 CTACATTAGAAGGATCTTCCCTCTGGAAATGTACCAACATCTTAGAGCTGAGGAAATA 3840
Qy 3841 AGAATTCAGACATAGTCCAGTCCAGACACAAATACCTTGGCCACAGACCCCATGAG 3900
Db 3841 AGAATTCAGACATAGTCCAGTCCAGACACAAATACCTTGGCCACAGACCCCATGAG 3900
Qy 3901 TGGGCGCGCTTCTCTTCTGACAGCAACGCGGGGCTCTTTAAATCAAGTCCACTGT 3960
Db 3901 TGGGCGCGCTTCTCTTCTGACAGCAACGCGGGGCTCTTTAAATCAAGTCCACTGT 3960
Qy 3961 GGTGTGAGGACCTTGTCAAGAACTCTGAGTGGTTCGGGGACAGGTGACCACTGCT 4020
Db 3961 GGTGTGAGGACCTTGTCAAGAACTCTGAGTGGTTCGGGGACAGGTGACCACTGCT 4020

Db	6181	GGGCTTCACTCGACCACTCCGCTACCGTCAAGTTGGGCCCTGATTGACCGACAGATCTT	6240
Qy	6241	CCGCTTCACTGAGGAAGGCATGGTCAACGCCCGTTTTTGACTACAACTATGACAAACAGCTT	6300
Db	6241	CCGCTTCACTGAGGAAGGCATGGTCAACGCCCGTTTTTGACTACAACTATGACAAACAGCTT	6300
Qy	6301	CCGGTGACACAGCATGACGGCTGTGATCAACAGACCCCACTGCCATTGATCTCTATCG	6360
Db	6301	CCGGTGACACAGCATGACGGCTGTGATCAACAGACCCCACTGCCATTGATCTCTATCG	6360
Qy	6361	CTATGATGATGTGTACGGCAAGACAGACAGTGTGGGAAGTTTTGGTGTCTCAATTACATGA	6420
Db	6361	CTATGATGATGTGTACGGCAAGACAGACAGTGTGGGAAGTTTTGGTGTCTCAATTACATGA	6420
Qy	6421	CATTAAACAGATCATCAACACAGCTGTCAATCAACCCACCAAGCATTTTGATGCATATGG	6480
Db	6421	CATTAAACAGATCATCAACACAGCTGTCAATCAACCCACCAAGCATTTTGATGCATATGG	6480
Qy	6481	CAGGATGAAGGAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGAGTACCGCTCCA	6540
Db	6481	CAGGATGAAGGAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGAGTACCGCTCCA	6540
Qy	6541	GTATGATAACTGGGGCGAGTAGTGAAGAAGGAGCTGAAGGTAGGACCCCTACGCCCAATAC	6600
Db	6541	GTATGATAACTGGGGCGAGTAGTGAAGAAGGAGCTGAAGGTAGGACCCCTACGCCCAATAC	6600
Qy	6601	CATTCGGTACTCTATGAGTATGATGTGACGGCAGCTCGACAGTCTCCCATCAATGA	6660
Db	6601	CATTCGGTACTCTATGAGTATGATGTGACGGCAGCTCGACAGTCTCCCATCAATGA	6660
Qy	6661	CAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGG	6720
Db	6661	CAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCACCTTACTGAGCCCTGG	6720
Qy	6721	GAACAGTGCACGGCTACACACATACGGTATGACATCCGGACCGCATCACTCGGTGGG	6780
Db	6721	GAACAGTGCACGGCTACACACATACGGTATGACATCCGGACCGCATCACTCGGTGGG	6780
Qy	6781	TGACGTGCAATACAAGATGATGAGGATGCTTCCCTGAGCAGCGGGCGGTGATATCTT	6840
Db	6781	TGACGTGCAATACAAGATGATGAGGATGCTTCCCTGAGCAGCGGGCGGTGATATCTT	6840
Qy	6841	TGAGTACAACCTCAGCTGGCTGCTCATCAAGGCGCTACAACCGGGCTGGCAGCTGGAGTGT	6900
Db	6841	TGAGTACAACCTCAGCTGGCTGCTCATCAAGGCGCTACAACCGGGCTGGCAGCTGGAGTGT	6900
Qy	6901	CAGGTACCGCTACGATGGCTTGGGGCGGCGGTGTCCAGCAAGAGCAGCCACAGCACCA	6960
Db	6901	CAGGTACCGCTACGATGGCTTGGGGCGGCGGTGTCCAGCAAGAGCAGCCACAGCACCA	6960
Qy	6961	CCTGCAGTTCTTATGACAGACCTGACCAACCCCAACCAAGGTCAACCACTGTACAAACCA	7020
Db	6961	CCTGCAGTTCTTATGACAGACCTGACCAACCCCAACCAAGGTCAACCACTGTACAAACCA	7020
Qy	7021	CTCCAGCTTGAGATCACTCCCTCTACTACGATTTGTGACAAATCGGACCCCTCTGCTGT	7080
Db	7021	CTCCAGCTTGAGATCACTCCCTCTACTACGATTTGTGACAAATCGGACCCCTCTGCTGT	7080
Qy	7081	GCTGAGCAGTGGTGAAGTATTTACATAGCTTTGTGACAAATCGGACCCCTCTGCTGT	7140
Db	7081	GCTGAGCAGTGGTGAAGTATTTACATAGCTTTGTGACAAATCGGACCCCTCTGCTGT	7140
Qy	7141	CTTTAGTGGAAACAGTTTGAATGATCAAGCAAAATCCCTGTACACAGCTATGGGAGATCTA	7200
Db	7141	CTTTAGTGGAAACAGTTTGAATGATCAAGCAAAATCCCTGTACACAGCTATGGGAGATCTA	7200
Qy	7201	CATTGGATACCAACCCCAACCTTCAGATCATCATAGCTACCATGGTGGCTCTATGATCC	7260
Db	7201	CATTGGATACCAACCCCAACCTTCAGATCATCATAGCTACCATGGTGGCTCTATGATCC	7260
Qy	7261	ACTCACCAAGCTGTGTCCACATGGCGCGGAGATTATGATGTGCTGCGCGGACGCTGGAC	7320

Db	7261	ACTCACCAAGCTTGTCCACATGGCGCCGGGAGATTATGATGTGCTGCCCGGACCGCTGGAC	7320
Qy	7321	TAGCCCCAGACCACGAGCTGTGGAGCACCCTTAGTAGCAGCAACGTCATGCTCTTTTAATCT	7380
Db	7321	TAGCCCCAGACCCGAGCTGTGGAGCACCCTTAGTAGCAGCAACGTCATGCTCTTTTAATCT	7380
Qy	7381	CTATATGTTCAAAAACAACAACCCCATCAGCAACTCCACGAGCATCAAGTGTTCATGAC	7440
Db	7381	CTATATGTTCAAAAACAACAACCCCATCAGCAACTCCACGAGCATCAAGTGTTCATGAC	7440
Qy	7441	AGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACAAAGTGATCCCTGGTTA	7500
Db	7441	AGATGTTAAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACAAAGTGATCCCTGGTTA	7500
Qy	7501	TCCCAAAACCAAGACATGGATGCCATGTGAACCCCTCTACGAGCTCATCCACACAGATGAA	7560
Db	7501	TCCCAAAACCAAGACATGGATGCCATGTGAACCCCTCTACGAGCTCATCCACACAGATGAA	7560
Qy	7561	AACCCAGAGTGGGACAAACGCAAGTCTATCTCTGGGGTACAGTGTGAAGTACAGAAGCA	7620
Db	7561	AACCCAGAGTGGGACAAACGCAAGTCTATCTCTGGGGTACAGTGTGAAGTACAGAAGCA	7620
Qy	7621	GCTCAAGGCCCTTTGTCACTTTAGAACGGTTGACCAAGCTCTATGGGTCCACAATCACCAG	7680
Db	7621	GCTCAAGGCCCTTTGTCACTTTAGAACGGTTGACCAAGCTCTATGGGTCCACAATCACCAG	7680
Qy	7681	CTGCCAGAGGCTCCAAAGACCAAGAAAGTTTGATCCAGCGGTCTAGTCTTTGGCAAGGG	7740
Db	7681	CTGCCAGAGGCTCCAAAGACCAAGAAAGTTTGATCCAGCGGTCTAGTCTTTGGCAAGGG	7740
Qy	7741	GGTCAAGTTTGCCTTTGAAGGATGCCGAGTGACACAGACATCATCAGTGTGCCCAATGA	7800
Db	7741	GGTCAAGTTTGCCTTTGAAGGATGCCGAGTGACACAGACATCATCAGTGTGCCCAATGA	7800
Qy	7801	GGATGGCGAAGGGTTCTGCCATCTTTGAACCAATGCCACTACTAGAGAACTGTGCACATT	7860
Db	7801	GGATGGCGAAGGGTTCTGCCATCTTTGAACCAATGCCACTACTAGAGAACTGTGCACATT	7860
Qy	7861	CACCATTTAGTGGGTGTGATACCACTACTTTGTGAACACGAGCACTTCAGAGGTGACCT	7920
Db	7861	CACCATTTAGTGGGTGTGATACCACTACTTTGTGAACACGAGCACTTCAGAGGTGACCT	7920
Qy	7921	GGCCATCCTGGGCTCAGTGGGGGGCGGAAACCCCTGGAGAAATGGGTCAACGTCACTGT	7980
Db	7921	GGCCATCCTGGGCTCAGTGGGGGGCGGAAACCCCTGGAGAAATGGGTCAACGTCACTGT	7980
Qy	7981	GTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGTACACAGACATCAGCTCCA	8040
Db	7981	GTCCAGATCAACACAGTACTTAATGGCAGGACTAGACGTACACAGACATCAGCTCCA	8040
Qy	8041	GTACGGGGCACTGTGCTTTGAACACACGCTACGGGCAACCGTTGGATGAGAGAAAGGCAG	8100
Db	8041	GTACGGGGCACTGTGCTTTGAACACACGCTACGGGCAACCGTTGGATGAGAGAAAGGCAG	8100
Qy	8101	GGTCTCTGGAGTGGCCCGGACAGAGCGGTGGCCCAAGCGTTGGGCCCGCGAGCAGAG	8160
Db	8101	GGTCTCTGGAGTGGCCCGGACAGAGCGGTGGCCCAAGCGTTGGGCCCGCGAGCAGAG	8160
Qy	8161	ACTCGGGAAGGGAGGAAGCGCTCGGGGCTCGACAGAGGGGGAAGCAGCAGGTGCT	8220
Db	8161	ACTCGGGAAGGGAGGAAGCGCTCGGGGCTCGACAGAGGGGGAAGCAGCAGGTGCT	8220
Qy	8221	GAGCACAGGCGGTGCAAGGCTACAGCGCTTTTTCGTGATCTCTGTGAGCAGTACCC	8280
Db	8221	GAGCACAGGCGGTGCAAGGCTACAGCGCTTTTTCGTGATCTCTGTGAGCAGTACCC	8280
Qy	8281	AGAACTCTCAGACAGCGCCCAACACTCCACTTCATGACACAGCGAGTGGCGGGAG	8340
Db	8281	AGAACTCTCAGACAGCGCCCAACACTCCACTTCATGACACAGCGAGTGGCGGGAG	8340
Qy	8341	GTGACAGAGGGAC	8354
Db	8341	GTGACAGAGGGAC	8354

QY 1535 AGGCTCTTAACCCAGGAGCGCGGAGCCTAGAGGGACCCCGCCAGCTCTCGGGAACT 1594
Db 1501 AGGCTCTTAACCCAGGAGCGCGGAGCCTAGAGGGACCCCGCCAGCTCTCGGGAACT 1560
QY 1595 GTGCCCCCTCCAGCCATGACAGAGGCTTCATCCAGTATTGGATTTCAGGAATCTGGCAC 1654
Db 1561 GTGCCCCCTCCAGCCATGACAGAGGCTTCATCCAGTATTGGATTTCAGGAATCTGGCAC 1620
QY 1655 TTGGCTTTTACAATGACGGAAGAGTACAGAGTGGTTCTTCTTCACCATGCCCATT 1714
Db 1621 TTGGCTTTTACAATGACGGAAGAGTACAGAGTGGTTCTTCTTCCTCTCTCCCC-----A 1674
QY 1715 GAGTCGTGGTAATAGTCCCCAGCAACTGCTATGGCAATGGTACTGCATCTCTGGGACC 1774
Db 1675 GAGTCGTGGTAATAGTCCCCAGCAACTGCTATGGCAATGGTACTGCATCTCTGGGACC 1734
QY 1775 TGCACATGCTTCTCGGTGTTCTGGGCCCGGACGTGTGGCAGAGCCTCTGCCCGGCTGCTC 1834
Db 1735 TGCACATGCTTCTCGGTGTTCTGGGCCCGGACGTGTGGCAGAGCCTCTGCCCGGCTGCTC 1794
QY 1835 TGTAGCGGAATGSCCAATACATAAGAGGAGATGCTTGTGCCACAGTGGCTCGGAAGGC 1894
Db 1795 TGTAGCGGAATGSCCAATACATAAGAGGAGATGCTTGTGCCACAGTGGCTCGGAAGGC 1854
QY 1895 GCTGAGTCGATGTGCCCAACCAACCAAGTGTATCGATGTGSCCTGCGACCAACCATGGCACC 1954
Db 1855 GCTGAGTCGATGTGCCCAACCAACCAAGTGTATCGATGTGSCCTGCGACCAACCATGGCACC 1914
QY 1955 TGCATACGGGACCTGCATCTGCAACCCCTGGCTACAAAGGCGAGAGCTGTGAGGAAGTG 2014
Db 1915 TGCATACGGGACCTGCATCTGCAACCCCTGGCTACAAAGGCGAGAGCTGTGAGGAAGTG 1971
QY 2015 GACTGCAATGACCCCAATGTTAGGCGGGGTGTCTGCGTGAGAGCGGAATGCCATTGC 2074
Db 1972 GACTGCAATGACCCCAATGTTAGGCGGGGTGTCTGCGTGAGAGCGGAATGCCATTGC 2031
QY 2075 TTTGTGGATGGGAGGACCACTGCGAGACCCCGAGGCCACATGCTTAGACCAAGTGT 2134
Db 2032 TCTGTGGATGGGAGGACCACTGCGAGACCCCGAGGCCACATGCTTAGACCAAGTGT 2091
QY 2135 TCAGGCCACCGAACCTTCTCCCGGACACCGGCTTTGACGTGTGACCCAGCTGGACT 2194
Db 2092 TCAGGCCACCGAACCTTCTCCCGGACACCGGCTTTGACGTGTGACCCAGCTGGACT 2151
QY 2195 GGCACGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGGTGCCATGSCGTGTGCGTA 2254
Db 2152 GGCACGACTGTTCTATCGAGATCTGTGCTGCCGACTGTGGTGCCATGSCGTGTGCGTA 2211
QY 2255 GGGGGACCTGCGCTGCGAGGATGCTGATGGGGGACCTGCGACACAGCGGGCTGCG 2314
Db 2212 GGGGGACCTGCGCTGCGAGGATGCTGATGGGGGACCTGCGACACAGCGGGCTGCG 2271
QY 2315 CACCCGGCTGTGCGGAGCATGGGACCTGCGCGGACGGCAAGTGGAGTGCAGCCCTGGC 2374
Db 2272 CACCCGGCTGTGCGGAGCATGGGACCTGCGCGGACGGCAAGTGGAGTGCAGCCCTGGC 2331
QY 2375 TGGATGGCGAAACACTGCACCATC-----GCTCACTATCTGATAGGTAGTTAA---- 2424
Db 2332 TGGATGGCGAAACACTGCACCATCTCCCTAGCTCACTATCTGATAGGTAGTTAACTT 2391
QY 2425 --AGAGGTTGCCCTGGGTTGTCAATGGCAAGGAGATGCTTACCTTAGACCTGATGGT 2482
Db 2392 TCAGAGGTTGCCCTGGGTTGTCAATGGCAAGGAGATGCTTACCTTAGACCTGATGGT 2451
QY 2483 TGGCATGCTGTGCCAGTGGCTGAGAGGAGCTGGCTGTGACATTTCCATGGAGACT 2542
Db 2452 TGGCATGCTGTGCCAGTGGCTGAGAGGAGCTGGCTGTGACATTTCCATGGAGACT 2511
QY 2543 GCCTGGGTCACAGCAAGACATGATGGAGATGGCTGGTGGATGCAATGACACCTGAC 2602
Db 2512 GCCTGGGTCACAGCAAGACATGATGGAGATGGCTGGTGGATGCAATGACACCTGAC 2571

QY 2603 TGCTGCTCCAGCCCTGTGCCATATCAACCCGCTGTGCTTGGCTTCCCTAACCCCTG 2662
Db 2572 TGCTGCTCCAGCCCTGTGCCATATCAACCCGCTGTGCTTGGCTTCCCTAACCCCTG 2631
QY 2663 GACATCATCCAGGAGACAGGTCCCTGTGTCAAGCAAACTACACTCTCTTATGAC 2722
Db 2632 GACATCATCCAGGAGACAGGTCCCTGTGTCAAGCAAACTACACTCTCTTATGAC 2691
QY 2723 CGCATCAAGTTCTCGTGGGAGGACAGCACGCAATAATCCCGGGAGAACCCCTTT 2782
Db 2692 CGCATCAAGTTCTCGTGGGAGGACAGCACGCAATAATCCCGGGAGAACCCCTTT 2751
QY 2783 GATGAGGCGCATGCTGTGTATTCTGTGCAAGTGTACATCAGATGGAACCCCTG 2842
Db 2752 GATGAGGCGCATGCTGTGTATTCTGTGCAAGTGTACATCAGATGGAACCCCTG 2811
QY 2843 GTTGTGTGAACATCAGTTTTCTCAATAACCCCTCTCTTTGGATATACAAATCAGAGGCAA 2902
Db 2812 GTTGTGTGAACATCAGTTTTCTCAATAACCCCTCTCTTTGGATATACAAATCAGAGGCAA 2871
QY 2903 GATGCGAGTTTGCATTTGGTGAATAATGGCGCATCTCAATCCTTGGGTTCCAGCG 2962
Db 2872 GATGCGAGTTTGCATTTGGTGAATAATGGCGCATCTCAATCCTTGGGTTCCAGCG 2931
QY 2963 GCACCTTTTCATCACAGAGCACACCCCTGTGGCTGCCATGGATCGCTTCTTCTCATG 3022
Db 2932 GCACCTTTTCATCACAGAGCACACCCCTGTGGCTGCCATGGATCGCTTCTTCTCATG 2991
QY 3023 GAAACATCATCATGAGCATGAGGAAATGAGATTCCAGCTGTGACTGAGCAATTTT 3082
Db 2992 GAAACATCATCATGAGCATGAGGAAATGAGATTCCAGCTGTGACTGAGCAATTTT 3051
QY 3083 GCCCGCCCAACCCAGTGTCTCCATCCCACTCCAGCTGCTTCCGAGCTCTGTGCA 3142
Db 3052 GCCCGCCCAACCCAGTGTCTCCATCCCACTCCAGCTGCTTCCGAGCTCTGTGCA 3111
QY 3143 GAGAAGGCCCATTTGTCCGAAATTCAGGCTTTGAGGAGAAATCTCTATCTTGGC 3202
Db 3112 GAGAAGGCCCATTTGTCCGAAATTCAGGCTTTGAGGAGAAATCTCTATCTTGGC 3171
QY 3203 TGCAGATGAGCTGAGCTACTGAGCAGCGGACCCCTGCGTACAAATCTGCTCAGG 3262
Db 3172 TGCAGATGAGCTGAGCTACTGAGCAGCGGACCCCTGCGTACAAATCTGCTCAGG 3231
QY 3263 ATCAGCTCACCACCCAGCATCCCTTCAACCTCATGAAGGTGCACCTCATGTAGCG 3322
Db 3232 ATCAGCTCACCACCCAGCATCCCTTCAACCTCATGAAGGTGCACCTCATGTAGCG 3291
QY 3323 GTGAGGCGCGCTCTTCAGGAAGTGTGCTGAGCCCGAGACCTCTCTATTATTTC 3382
Db 3292 GTGAGGCGCGCTCTTCAGGAAGTGTGCTGAGCCCGAGACCTCTCTATTATTTC 3351
QY 3383 ATTTGGGACAAGACAGAGCTCTACAACAGAGGTGTTGGGCTTTCAGAGCCCTTGT 3442
Db 3352 ATTTGGGACAAGACAGAGCTCTACAACAGAGGTGTTGGGCTTTCAGAGCCCTTGT 3411
QY 3443 TCCGTGGGTTATGAATATGAATCTTGGCCAGATCTTAATCTGTGGGAAAAAGAACACA 3502
Db 3412 TCCGTGGGTTATGAATATGAATCTTGGCCAGATCTTAATCTGTGGGAAAAAGAACACA 3471
QY 3503 GTGCTGAGGCGCTGAGAAATGAGCGCTTCAAGCTTGGAGATGAGGCTAGACAAACAT 3562
Db 3472 GTGCTGAGGCGCTGAGAAATGAGCGCTTCAAGCTTGGAGATGAGGCTAGACAAACAT 3531
QY 3563 CATGCCCTCAACATTTCAAAGTGGT---ATCCTGCACAAAGGAATGGGGAAACAGTTT 3619
Db 3532 CATGCCCTCAACATTTCAAAGTGGTGGCATCTGCAAAAGGAATGGGGAAACAGTTT 3591
QY 3620 GTGTCTCAGAGCCCTCTGTCTATTGGGAGCATATGAGGCAATGGGCGCGGAGAGCATC 3679
Db 3592 GTGTCTCAGAGCCCTCTGTCTATTGGGAGCATATGAGGCAATGGGCGCGGAGAGCATC 3651
QY 3680 TCCTGCCCCAGCTGCAACGGGCTTGTGTGACGGCAACAAGCTCCTTGGCCCCCAGTGGCCCTC 3739

Db	3652	TCCTGCCCAAGCTGCAAGCGCTTGTGACGGCAACAAGCTCCTGCGCCCAAGTGGCCCTC	3711
Qy	3740	ACCTGTGGCTCTGACGGGAGCTCTATGTGGGTGATTTCACTACATTAGAGGATCTTC	3799
Db	3712	ACCTGTGGCTCTGACGGGAGCTCTATGTGGGTGATTTCACTACATTAGAGGATCTTC	3771
Qy	3800	CCCTCTGGAAATGTCACCAACATCTTAGAGCTCAGG-----AATAAAGATTTCAGACAT	3853
Db	3772	CCCTCTGGAAATGTCACCAACATCTTAGAGCTCAGGTCAGAAATAAAGATTTCAGACAT	3831
Qy	3854	AGTCACAGTCACGACACAAATACTACTGGGCCACAGACCCCATAGTGGGCGCTTCCTC	3913
Db	3832	AGTCACAGTCACGACACAAATACTACTGGGCCACAGACCCCATAGTGGGCGCTTCCTC	3891
Qy	3914	CTTTCTGACAGCAACAGCGCGGCTCTTAATAATCAAGTCCACCTGCTGGTGGAGGAC	3973
Db	3892	CTTTCTGACAGCAACAGCGCGGCTCTTTAAATCAAGTCCACCTGCTGGTGGTGAAGGAC	3951
Qy	3974	CTTTCTCAAGAACTCTGAGGTGGTTGGGGGACAGGTGACCAAGTGCCTCCCTTTGATGAC	4033
Db	3952	CTTTCTCAAGAACTCTGAGGTGGTTGGGGGACAGGTGACCAAGTGCCTCCCTTTGATGAC	4011
Qy	4034	ACTCGCTGGGGGATGTGGGAAGGCCACAGAGCCACACTCACCAATCCAGGGGT---	4090
Db	4012	ACTCGCTGGGGGATGTGGGAAGGCCACAGAGCCACACTCACCAATCCAGGGGTCC	4071
Qy	4091	-----ATTACAGTGGACAAAGTTTGGGCTGATCTACTTCTGGATGGCACCATGATCAGA	4144
Db	4072	CCAGGCATTAACAGTGGACAAAGTTTGGGCTGATCTACTTCTGGATGGCACCATGATCAGA	4131
Qy	4145	CGCATCGATCAGAAATGGATCATCTCCACCTGCTGGCTCTTAATGATCTCATATGACC	4204
Db	4132	CGCATCGATCAGAAATGGATCATCTCCACCTGCTGGCTCTTAATGATCTCATATGACC	4191
Qy	4205	CGGCACCTCAGCTGATTTCTGATGATATTTCCCGAGTAA-----CTGGAG	4255
Db	4192	CGGCACCTCAGCTGATTTCTGATGATATTTCCCGAGTAAAGCAGGTTTACCTGGAG	4251
Qy	4256	TGGCCACAGACTTAGCCATCAACCCAAATGGACAACTCATTTATGCTCTCGAACAAAT	4315
Db	4252	TGGCCACAGACTTAGCCATCAACCCAAATGGACAACTCATTTATGCTCTCGAACAAAT	4311
Qy	4316	GTGCTCTGCAATCTCTGAACCCAGGTGGGATTTGCGCGGGAGGCCATGGAC	4375
Db	4312	GTGCTCTGCAATCTCTGAACCCAGGTGGGATTTGCGCGGGAGGCCATGGAC	4371
Qy	4376	TGCCAGGTCCCTGGCAATGACCACTTCTCTGAAGCAAGGTGGCCATCCACGCAACCTCG	4435
Db	4372	TGCCAGGTCCCTGGCAATGACCACTTCTCTGAAGCAAGGTGGCCATCCACGCAACCTCG	4431
Qy	4436	GAGTCAGCCACCGTTTGGCTGTTTCAACAATGGGGTCTGTATATGCTGAGACTGAT	4495
Db	4432	GAGTCAGCCACCGTTTGGCTGTTTCAACAATGGGGTCTGTATATGCTGAGACTGAT	4491
Qy	4496	GAGAAAAGATCAACCGCATCAGCAGGTCAACACTAGTGGAGAGATCTCACTCGTTGCT	4555
Db	4492	GAGAAAAGATCAACCGCATCAGCAGGTCAACACTAGTGGAGAGATCTCACTCGTTGCT	4551
Qy	4556	GGGGCCCCAGTGGCTGTGATGTAAATAATGATGCCAACTGTGATTTTCTGGAGAC	4615
Db	4552	GGGGCCCCAGTGGCTGTGATGTAAATAATGATGCCAACTGTGATTTTCTGGAGAC	4611
Qy	4616	GATGGTTATGCCAAGGATGAAAGTTAAATACCCATCTTCTTGGCTGTGCTGTGAT	4675
Db	4612	GATGGTTATGCCAAGGATGAAAGTTAAATACCCATCTTCTTGGCTGTGCTGTGAT	4671
Qy	4676	GGGAGGCTCTACGTGGCGGACTTTGGGAACATCGAAATTCGGTTATCCGGAAGAACAG	4735
Db	4672	GGGAGGCTCTACGTGGCGGACTTTGGGAACATCGAAATTCGGTTATCCGGAAGAACAG	4731
Qy	4736	CTTTCTCACAACCCAGAACATGATGAGTGTCTTACCAATTTGACAGGAGCTCTAT	4795

Db	4732	CTTTCTCACAACCCAGAACATGATGAGTGTCTTCAACAATGACCAAGGAGCTCTAT	4791
Qy	4796	CTGTTTGTATACCAACCGCAAGCACTGTACACCAAGAGCTGCCCCACAGAGACTACCTG	4855
Db	4792	CTGTTTGTATACCAACCGCAAGCACTGTACACCAAGAGCTGCCCCACAGAGACTACCTG	4851
Qy	4856	TACAACTTCACTACCTGCGGACCGGCAATCACTACTCATCAAGCAACAATGGCAAC	4915
Db	4852	TACAACTTCACTACCTGCGGACCGGCAATCACTACTCATCAAGCAACAATGGCAAC	4911
Qy	4916	ATGCTAATATGTCGCCAGAGACTCTACTGGGATGCCCTCTGGCTGTGTGCTCCAGATGGC	4975
Db	4912	ATGCTAATATGTCGCCAGAGACTCTACTGGGATGCCCTCTGGCTGTGTGCTCCAGATGGC	4971
Qy	4976	CAGGTGTACTGGGTGACCATGGGCACCAACAGTGCATCTCAAGAGTGTGACCAACAAGGA	5035
Db	4972	CAGGTGTACTGGGTGACCATGGGCACCAACAGTGCATCTCAAGAGTGTGACCAACAAGGA	5031
Qy	5036	CACAGTTTGGCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAAT	5095
Db	5032	CACAGTTTGGCATGATGACATACCATGGCAATTCGGGCTTCTGGCAACCAAAAGCAAT	5091
Qy	5096	GAAAACGATGACCAACAATTTATGATGACGACAGCTTTTGGCGCTTGACAAATGTGACC	5155
Db	5092	GAAAACGATGACCAACAATTTATGATGACGACAGCTTTTGGCGCTTGACAAATGTGACC	5151
Qy	5156	TTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGTATACAGACAGTTTCACTGATGTCAG	5215
Db	5152	TTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTGTATACAGACAGTTTCACTGATGTCAG	5211
Qy	5216	GTAGAGACCTCCAGCAAGGATGTCCACATACCAACCAACCTGCTGCTCAGGGGCC	5275
Db	5212	GTAGAGACCTCCAGCAAGGATGTCCACATACCAACCAACCTGCTGCTCAGGGGCC	5271
Qy	5276	TTCTACACACTGCTGCAAGCAACCAAGTCCGGAACAGTACTACATCGGGGCGGATGGTCC	5335
Db	5272	TTCTAC-----GACCAAGTCCGGAACAGTACTACATCGGGGCGGATGGTCC	5319
Qy	5336	TTGGGCTGTGCTGGCCACCGCATGAGGTGGCGCTGAGACTGAGCCCACTTCTGCTG	5395
Db	5320	TTGGGCTGTGCTGGCCACCGCATGAGGTGGCGCTGAGACTGAGCCCACTTCTGCTG	5379
Qy	5396	GCTGSCACCGTCAACCCCAACCGTGGGCAAGAGGAATGTACGCTGCCATCGCAACCGC	5455
Db	5380	GCTGSCACCGTCAACCCCAACCGTGGGCAAGAGGAATGTACGCTGCCATCGCAACCGC	5439
Qy	5456	CTCAACCTGTGAGTGGCGCCAGCGCAAGAGAGAGGCTCGGGGCGCAGGCTCTGCTTT	5515
Db	5440	CTCAACCTGTGAGTGGCGCCAGCGCAAGAGAGAGGCTCGGGGCGCAGGCTCTGCTTT	5499
Qy	5516	GGGCGCGGCTCGGGGTG-----CACAAACGAAATCTCTATCTCTGGACTTTGAT	5566
Db	5500	GGGCGCGGCTCGGGGTGCTCCAGGTTCAACAACGAAATCTCTATCTCTGGACTTTGAT	5559
Qy	5567	CGCGTAAACACGACAGAGAGATCTATGATGACACCGCAAGTTCACTTCTGAGTCTG	5626
Db	5560	CGCGTAAACACGACAGAGAGATCTATGATGACACCGCAAGTTCACTTCTGAGTCTG	5619
Qy	5627	TACGACAGCGGGGCGCCAGGCTCTGTGTCACCCAGCAGGCTGAAATGGTGTCAAC	5686
Db	5620	TACGACAGCGGGGCGCCAGGCTCTGTGTCACCCAGCAGGCTGAAATGGTGTCAAC	5679
Qy	5687	GTGACATACTCCCTCGGGGTTCATTGCTGGCATCCAGAGGGGCATGCTCTGAAAGA	5746
Db	5680	GTGACATACTCCCTCGGGGTTCATTGCTGGCATCCAGAGGGGCATGCTCTGAAAGA	5739
Qy	5747	ATGGAATACGACAGCGGGGCGGCATCACATCCAGGATCTTCGCTGATGGGAACATGG	5806
Db	5740	ATGGAATACGACAGCGGGGCGGCATCACATCCAGGATCTTCGCTGATGGGAACATGG	5799
Qy	5807	AGCTACATACTTACGAGAGTCCATGCTGTCTACTACACGCCAGAGCAGTATATC	5866
Db	5800	AGCTACATACTTACGAGAGGCGAGGTG-----	5827

QY 5867 TTTGAGTTTCAGAAAGATGACCGCTCTCTTCTGTGTAGCATGCCCCAACGTCGGCGCGCAG 5926
 Db 5928 -TCAGATTTCAGAAAGATGACCGCTCTCTTCTGTGTAGCATGCCCCAACGTCGGCGCGCAG 5986
 QY 5927 ACATAGAGACCATCCGCTCAGTGGGTACTACAGAAACATCTATCAGCCCCCTGAGGGC 5986
 Db 5987 ACATAGAGACCATCCGCTCAGTGGGTACTACAGAAACATCTATCAGCCCCCTGAGGGC 5946
 QY 5987 AATGCCCTCAGTCATACAGGACTTCACCTGAGGATGGGCACCTCTTACACCTTCTACCTG 6046
 Db 5947 AATGCCCTCAGTCATACAGGACTTCACCTGAGGATGGGCACCTCTTACACCTTCTACCTG 6006
 QY 6047 GGCAGTCGGCGCAGGATGATATACAGTATGGCAACTGTCAAGCTGGGAGAGAGCTC 6106
 Db 6007 GGCAGTCGGCGCAGGATGATATACAGTATGGCAACTGTCAAGCTGGGAGAGAGCTC 6066
 QY 6107 TATGACACCAACAGGTCAGTTTCACTATGACGAGACGGCAGGCATGCTGGAAGACCATC 6166
 Db 6067 TATGACACCAACAGGTCAGTTTCACTATGACGAGACGGCAGGCATGCTGGAAGACCATC 6126
 QY 6167 AACCTACAGATGAGGCTTCACTGACCATCCGCTACCGTCAGATTTGGGCCCTTGATT 6226
 Db 6127 AACCTACAGATGAGGCTTCACTGACCATCCGCTACCGTCAGATTTGGGCCCTTGATT 6186
 QY 6227 GACCGACAGATCTTCGCTTCACTGAGGAAGGATGTCACACCGCTTTTGTACTACAAC 6286
 Db 6187 GACCGACAGATCTTCGCTTCACTGAGGAAGGATGTCACACCGCTTTTGTACTACAAC 6246
 QY 6287 TATGACAAAGCTTCGGGTGACGACGATGACGATGAGTCAACGAGACCCCACTGCC 6346
 Db 6247 TATGACAAAGCTTCGGGTGACGACGATGAGTCAACGAGACCCCACTGCC 6306
 QY 6347 ATTGATCTCATCGCTATGATGATGTCAGGCAAGCAGAGCAGTTTGGGAGTTTGGT 6406
 Db 6307 ATTGATCTCATCGCTATGATGATGTCAGGCAAGCAGAGCAGTTTGGGAGTTTGGT 6366
 QY 6407 GTCAATTACTATGACATTAACAGATCATCACACAGCTGTCATGACCCACACCAAGCAT 6466
 Db 6367 GTCAATTACTATGACATTAACAGATCATCACACAGCTGTCATGACCCACACCAAGCAT 6426
 QY 6467 TTTGATGATATGCGAGGATGAAGGAGTGCAGTATGAGATCTTCGCTCGCTCATGTAC 6526
 Db 6427 TTTGATGATATGCGAGGATGAAGGAGTGCAGTATGAGATCTTCGCTCGCTCATGTAC 6486
 QY 6527 TGGATGACCTCCAGTATGATAACATGGGCGGAGTGTGAAGAGGAGCTGAAGGTAGGA 6586
 Db 6487 TGGATGACCTCCAGTATGATAACATGGGCGGAGTGTGAAGAGGAGCTGAAGGTAGGA 6546
 QY 6587 CCCTAGCCCAATACCACTCGCTACTCTATGATGATGCTGACCGCCAGCTGCGAGACA 6646
 Db 6547 CCCTAGCCCAATACCACTCGCTACTCTATGATGATGCTGACCGCCAGCTGCGAGACA 6606
 QY 6647 GTCTCCATCAATGACAGGACCTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCAC 6706
 Db 6607 GTCTCCATCAATGACAGGACCTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCAC 6666
 QY 6707 TTAAGTCCCTGGGAAACAGTGCACGGCTCACACCACTACGGTATGACATCCCGGACCGC 6766
 Db 6667 TTAAGTCCCTGGGAAACAGTGCACGGCTCACACCACTACGGTATGACATCCCGGACCGC 6726
 QY 6767 ATCACTCGGCTGGGTGAGTGCATACAGATGATGAGATGAGTTCCTGAGGAGCGG 6826
 Db 6727 ATCACTCGGCTGGGTGAGTGCATACAGATGATGAGATGAGTTCCTGAGGAGCGG 6786
 QY 6827 GGGGGTATATCTTTGAGTCAACTCAGCTGGGCTGCTCATCAAGGCTTACACCGGGCT 6886
 Db 6787 GGGGGTATATCTTTGAGTCAACTCAGCTGGGCTGCTCATCAAGGCTTACACCGGGCT 6846
 QY 6887 GGCAGCTGAGTGTACGATACCGCTACGATGGGCGGCGCTGTCCAGCAAGAGC 6946
 Db 6847 GGCAGCTGAGTGTACGATACCGCTACGATGGGCGGCGCTGTCCAGCAAGAGC 6906

QY 6947 AGCCACAGCCACCACTGAGTTCCTTATGAGACCTGACCAACCCCAAGGTCACC 7006
 Db 6907 AGCCACAGCCACCACTGAGTTCCTTATGAGACCTGACCAACCCCAAGGTCACC 6966
 QY 7007 CACCTGTACAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGATTCGCAAGGACAC 7066
 Db 6967 CACCTGTACAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGATTCGCAAGGACAC 7026
 QY 7067 CTCTTTGGCATGAGCTGAGCAGTGGTGTATGAGTTTATACATAGCTTGTGACAACTCGG 7126
 Db 7027 CTCTTTGGCATGAGCTGAGCAGTGGTGTATGAGTTTATACATAGCTTGTGACAACTCGG 7086
 QY 7127 ACCCTCTCTGCTCTTTAGTGGACACAGTTTGTATGATCAAGCAATCTGTACACAGC 7186
 Db 7087 ACCCTCTCTGCTCTTTAGTGGACACAGTTTGTATGATCAAGCAATCTGTACACAGC 7146
 QY 7187 TATGGGAGATCTACATGGATACCAACCCCACTTTCAGATCATCATAGGCTACCATGGT 7246
 Db 7147 TATGGGAGATCTACATGGATACCAACCCCACTTTCAGATCATCATAGGCTACCATGGT 7206
 QY 7247 GGCCTCTATGATCCTACTCACCAGCTTGTCCATGGCCGGCGGAGATATATGTGTCTG 7306
 Db 7207 GGCCTCTATGATCCTACTCACCAGCTTGTCCATGGCCGGCGGAGATATATGTGTCTG 7266
 QY 7307 GCGGACCTGACCTAGCCAGACCAAGCTGTGGAAGCACCTTTAGTAGCAGCAACGTC 7366
 Db 7267 GCGGACCTGACCTAGCCAGACCAAGCTGTGGAAGCACCTTTAGTAGCAGCAACGTC 7326
 QY 7367 ATGCCCTTTAATCTCTATATGTTCAAAAACCAACCCCATCAGCAACTCCAGGACATC 7426
 Db 7327 ATGCCCTTTAATCTCTATATGTTCAAAAACCAACCCCATCAGCAACTCCAGGACATC 7386
 QY 7427 AAGTGCCTTCATGACAGATGTTAAAGCTGGCTGCTACCTTTGGATTCAGCTACACAC 7486
 Db 7387 AAGTGCCTTCATGACAGATGTTAAAGCTGGCTGCTACCTTTGGATTCAGCTACACAC 7446
 QY 7487 GTGATCCCTGTTTATCCCAACCAAGCATGGATGTCATGGAACCCCTCTACGAGCTCATC 7546
 Db 7447 GTGATCCCTGTTTATCCCAACCAAGCATGGATGTCATGGAACCCCTCTACGAGCTCATC 7506
 QY 7547 CACACACAGATGAAACCCAGAGTGGGACAAACAGCA----- 7583
 Db 7507 CACACACAGATGAAACCCAGAGTGGGACAAACAGCAAGGTAATTCCTGCAAGGCTGC 7566
 QY 7584 -AGTCTATCTCTCGGGGTACAGTGTGAAGTACAGAAAGCAGCTCAAGGCTTTGTACCTTA 7642
 Db 7567 CAGTCTATCTCTCGGGGTACAGTGTGAAGTACAGAAAGCAGCTCAAGGCTTTGTACCTTA 7626
 QY 7643 GAACGGTTTGACAGCTCTATGGCTCCAAATCAAGCTGCCAGAGGCTCCAAAGACC 7702
 Db 7627 GAACGGTTTGACAGCTCTATGGCTCCAAATCAAGCTGCCAGAGGCTCCAAAGACC 7686
 QY 7703 AAGAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGAT 7762
 Db 7687 AAGAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGTCAAGTTTGCCTTGAAGAT 7746
 QY 7763 GGCAGGTGACACACAGACATCATCAGTGTGGCCCAATGAGGATGGGCGAAGGTTGCTGCC 7822
 Db 7747 GGCAGGTGACACACAGACATCATCAGTGTGGCCCAATGAGGATGGGCGAAGGTTGCTGCC 7806
 QY 7823 ATCTTGACCAATGCCACTACTAGAGAACCTGACATTCACCATGTATGGGGTGGATACC 7882
 Db 7807 ATCTTGACCAATGCCACTACTAGAGAACCTGACATTCACCATGTATGGGGTGGATACC 7866
 QY 7883 CATTAATTGTGAAACACAGGACCTTCAGAAGTCACTGGGCTCCTGGGCTCAGTGGG 7942
 Db 7867 CATTAATTGTGAAACACAGGACCTTCAGAAGTCACTGGGCTCCTGGGCTCAGTGGG 7926
 QY 7943 GGGGGCGAAACCTTGGGAATGGGGTCAACCTGCTGTGTCCAGATCAACACAGTACTT 8002
 Db 7927 GGGGGCGAAACCTTGGGAATGGGGTCAACCTGCTGTGTCCAGATCAACACAGTACTT 7986
 QY 8003 AATGCGAGGACTAGACGCTACACAGACATCCAGCTCCAGTACGGGGGCTGTGCTTGAAC 8062

Db 7987 AATGGCAGGACTAGAGCTACACAGACATCCAGCTCCAGTACGGGCACTGTGCTTGAAC 8046
QY 8063 ACAGCTACGGACACAGTGGATGAGGAGGACAGGCTCTGAGCTGCGCGGCGAG 8122
Db 8047 ACAGCTACGGGACACAGTGGATGAGGAGGACAGGCTCTGAGCTGCGCGGCGAG 8106
QY 8123 AGAGCCGTGCGCCCAAGCGTGGGCGCGAGCAGACAGAGACTGCGGGAAGGGAGGAAGGC 8182
Db 8107 AGAGCCGTGCGCCCAAGCGTGGGCGCGAGCAGACAGAGACTGCGGGAAGGGAGGAAGGC 8166
QY 8183 CTGGGGCCCTGGACAGAGGGGGAGAGCAGCAGCTGCTGAGCAGACAGGGGGGGTGCAGAGC 8242
Db 8167 CTGGGGCCCTGGACAGAGGGGGAGAGCAGCAGCTGCTGAGCAGACAGGGGGGGTGCAGAGC 8226
QY 8243 TAGACCGGCTTTTCGTGATCTCTGTCGAGCAGTACCCAGAACTGTACAGACAGCGCAAC 8302
Db 8227 TAGACCGGCTTTTCGTGATCTCTGTCGAGCAGTACCCAGAACTGTACAGACAGCGCAAC 8286
QY 8303 AACATCCACTTATGACAGACAGAGAGATGGCGGAGGTGACAGAGAGAC 8354
Db 8287 AACATCCACTTATGACAGACAGAGAGATGGCGGAGGTGACAGAGAGAC 8338

RESULT 4

US-10-042-865-1
; Sequence 1, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zehusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Esha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Perence L
; APPLICANT: Grosse, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Maliyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glennda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE OF INVENTION: Using the Same
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042,865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18

; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 8438
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-865-1
Query Match 93.1%; Score 7779.2; DB 13; Length 8438;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 811; Conservative 0; Mismatches 203; Indels 100; Gaps 11;
QY 32 GCCATGACAGTGAAGGAGAGAAACCTTACCGCTCGCTGACCGCGGCGCGGAGCGCGAG 91
Db 1 GCCATGACAGTGAAGGAGAGAAACCTTACCGCTCGCTGACCGCGGCGCGGAGCGCGAG 60
QY 92 CGCGCTACACACAGCTCGTCCGCGGACAGCGAGAGAGGCAAGCCCGCAGAAATCGTAC 151
Db 61 CGCGCTACACACAGCTCGTCCGCGGACAGCGAGAGAGGCAAGCCCGCAGAAATCGTAC 120
QY 152 AGTCCAGCGAGACCCCTGAAGGCTTACGACAGAGACGCGCGCTAGCCTATGGAGCGCG 211
Db 121 AGTCCAGCGAGACCCCTGAAGGCTTACGACAGAGACGCGCGCTAGCCTATGGAGCGCG 180
QY 212 GTCAGGACATTTGCGCGAGGAGCGAGGAATTTCTGCCGACAGGTGCCAACTTCACC 271
Db 181 GTCAGGACATTTGCGCGAGGAGCGAGGAATTTCTGCCGACAGGTGCCAACTTCACC 240
QY 272 CTGCGGAGCTGGGGCTTGAAGAGTAACCGCCCTCACGGGACCTGTACCGGACAGAC 331
Db 241 CTGCGGAGCTGGGGCTTGAAGAGTAACCGCCCTCACGGGACCTGTACCGGACAGAC 300
QY 332 ATTGGCTGCCCAATGGGCTACTCCATGGGGCTGGCTCTGATCGGCACATGGAGGCT 391
Db 301 ATTGGCTGCCCAATGGGCTACTCCATGGGGCTGGCTCTGATCGGCACATGGAGGCT 360
QY 392 GACACGCTGTGTCTCCCTGAGACACCCCGTGGCTGTGTGGGGCGCGAGACACAGGTACAGG 451
Db 361 GACACGCTGTGTCTCCCTGAGACACCCCGTGGCTGTGTGGGGCGCGAGACACAGGTACAGG 420
QY 452 CGAGCTCTGCTGTGTGAGCGGGCCAAATTCATCTCACTCACCAGCAGCAGGACAT 511
Db 421 CGAGCTCTGCTGTGTGAGCGGGCCAAATTCATCTCACTCACCAGCAGCAGGACAT 480
QY 512 GAAACACTGAGACTGATCATTCGCGGCGGCTGTGAGAAACCAAGCGCGCTCCGAGCGCG 571
Db 481 GAAACACTGAGACTGATCATTCGCGGCGGCTGTGAGAAACCAAGCGCGCTCCGAGCGCG 540
QY 572 CGCGCGCGCTCTCGACGCGCCACACCCCAACAGCAACAGCGCGGCTCCATTAATCC 631
Db 541 CGCGCGCGCTCTCGACGCGCCACACCCCAACAGCAACAGCGCGGCTCCATTAATCC 600
QY 632 CTGAACCGGGCAACTTTCAGCGGAGAGCAACCCAGCGCGCGCGCGCGCGCGCGCGCTCG 691
Db 601 CTGAACCGGGCAACTTTCAGCGGAGAGCAACCCAGCGCGCGCGCGCGCGCGCGCGCTCG 660
QY 692 CTCTCCGAGAGCGCCCTTCCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCG 751
Db 661 CTCTCCGAGAGCGCCCTTCCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTCG 720
QY 752 CTGCTCAACAGCAACATCCCTTGGAGACAGGAACTTAGCAGAGAGCCTATTCCTAGGG 811
Db 721 CTGCTCAACAGCAACATCCCTTGGAGACAGGAACTTAGCAGAGAGCCTATTCCTAGGG 780
QY 812 ACATTGAGGACAACTCATTTGAGATGAGATTTCTCGCGCGCTCCCGCCATGATGGGCT 871
Db 781 ACATTGAGGACAACTCATTTGAGATGAGATTTCTCGCGCGCTCCCGCCATGATGGGCT 840
QY 872 TACAGTGAAGGCACTTCTCTTCAAGCCTGAGGAGCACTCCCGCGCTCTTCTGCAACACA 931
Db 841 TACAGTGAAGGCACTTCTCTTCAAGCCTGAGGAGCACTCCCGCGCTCTTCTGCAACACA 900
QY 932 TCACCAGGGTACCCACTGACGTCCAGCAGAGTGTACTCTCTCCGCGCGCGCGCGCGCTCC 991

Db 901 TCACGAGGTACCCACTGACGTCCTCAGACAGTGTACTCTCTCCGCCCCGACCCCTGCCC 960
 QY 992 CGCAGACCTTCGCCCCGCGCTTTAACTCAGAGAGCCCTCCAGTACTGTAACTGG 1051
 Db 961 CGCAGACCTTCGCGCTCGCGCGCTTTAACTCAGAGAGCCCTCCAGTACTGTAACTGG 1020
 QY 1052 AAGTGCAGCCTGAGCGGCATCGTCACTCAGCCACTCTGTGTCATCTCTGCTGGCATAC 1111
 Db 1021 AAGTGCAGCCTGAGCGGCATCGTCACTCAGCCACTCTGTGTCATCTCTGCTGGCATAC 1080
 QY 1112 TTGTGCGCATGACCTGTTTGGCTAAACTGGACCTGACGCGATGAGAGGCGCAGATG 1171
 Db 1081 TTGTGCGCATGACCTGTTTGGCTAAACTGGACCTGACGCGATGAGAGGCGCAGATG 1140
 QY 1172 -----TATGAGATCAGGAGGACACAGCCAGAGTGTGGCTGTGCGCAACCGACGTCCTC 1224
 Db 1141 ACGGATTTATGAGATCAGCGAGGACACAGCCAGAGTGTGGCTGTGCGCAACCGAGTCTC 1200
 QY 1225 CCTATACCTCAGGGGCGACTGGCTTAGAGACCCCTGACAGAAAGGCAAGAAACCCAC 1284
 Db 1201 CCTATACCTCAGGGGCGACTGGCTTAGAGACCCCTGACAGAAAGGCAAGAAACCCAC 1260
 QY 1285 AGAAGAAAGCCAGTAGTTCCTTTCCAGAGGACAGTTTCATAGATCTCGGAGATTGA 1344
 Db 1261 AGAAGAAAGCCAGTAGTTCCTTTCCAGAGGACAGTTTCATAGATCTCGGAGATTGA 1320
 QY 1345 TGTGGGAAGCGAGCCTCCAGAGATTCCTCTGCGACCTTTCTGGAGATCTCAAGTGT 1404
 Db 1321 TGTGGGAAGCGAGCCTCCAGAGATTCCTCTGCGACCTTTCTGGAGATCTCAAGTGT 1380
 QY 1405 CATAGACCATCTGTCATCTGAATTCATATGTCCTCGGAAAGCGACCTGTTGG 1464
 Db 1381 CATAGACCATCTGTCATCTGAATTCATATGTCCTCGGAAAGCGACCTGTTGG 1440
 QY 1465 CATTTATGGCAGAAAGCCCTCCCTTCACATACACAGTTTGACTTTGTGGAGTGT 1524
 Db 1441 CATTTATGGCAGAAAGCCCTCCCTTCACATACACAGTTTGACTTTGTGGAGTGT 1500
 QY 1525 GATGCGAGAGCTCTTAACCGAGAGGCGCGAGCCTTAGAGGACCCCGCGCAGTC 1584
 Db 1501 GATGCGAGAGCTCTTAACCGAGAGGCGCGAGCCTTAGAGGACCCCGCGCAGTC 1560
 QY 1585 TCGGGAACTGTGCCCCCTCCAGCCATGAGCAGGCTTCATCCAGTATTTGGATTTCAGG 1644
 Db 1561 TCGGGAACTGTGCCCCCTCCAGCCATGAGCAGGCTTCATCCAGTATTTGGATTTCAGG 1620
 QY 1645 AATCTGGCACTTGGCTTTTACATACGGAAGAGTACAGAGTGGTTTCTCTCTCAC 1704
 Db 1621 AATCTGGCACTTGGCTTTTACATACGGAAGAGTACAGAGTGGTTTCTCTCTCAC 1680
 QY 1705 CACTGCCATTGAGTCGTGGATTAACCTGCCAGCACTGCTATGGCAATGGTACTGCAT 1764
 Db 1681 CACTGCCATT-----CTTGATTCTGGGCTCTCTGTTGGGTGATGAGAAATGGCT 1731
 QY 1765 CTCTGGGACCTGCCACTGTTCTCTGGGTTCTCGGCCCCGACTGTGGCAGAGCTCTCTG 1824
 Db 1732 TTCTGGAACTGTGCAATTGTTTCCAGGATTTCTGGGTCCGATTTGTTCAAGAGCCGCTG 1791
 QY 1825 CCCCCTGCTCTGAGCGGAAATGGCCAAATACATGAAGGAGAGATGTTGTGCCACAGTGG 1884
 Db 1792 TCCAGTGTATGTAGTGGCAACCGGCGAGTACTCCAGGGCCGCTGCTGTGTTTCAGCGG 1851
 QY 1885 CTGGAAGGCGCTGAGTGGATGTGCCCAACCAACAGTGTATCGATGTGGCTCGAGCA 1944
 Db 1852 CTGGAAGGCAACGAGTGTATGTGCCGACTACCCAGTGTATGACCCACAGTGTGGGG 1911
 QY 1945 CMAATGCACTGATCAGCGGACCTGCAATCTGCAACCTGCTACAGGGCGAGAGCTG 2004
 Db 1912 TCGTGGGATTTGATCATGGGCTCTGCTTGTGCAACTCAGGATACAAAGGAGAAAGTTG 1971
 QY 2005 TGAGGAAGTGGACTGATGAGCCCAACATGTTTCAGCCCGGGGTCTCGGTGAGAGGCA 2064

Db 1972 TGAAGAAGCTGACTGTATAGACCCCTGGGTGTTTCTAATCATGTGTGTGTATCCAGGGGA 2031
 QY 2065 ATGCCAATTGCTTTGTGGGATGAGGAGGACCAACTGCGAGACCCCGAGGCGCAGATGTT 2124
 Db 2032 ATGTCACTGCACTCCAGGATGAGGAGTAGCAATTTGTAAATACCTGAAGACCATGTGTCC 2091
 QY 2125 AGACAGAGTGTTCAGGCGCAGGAACCTTCTCCCGGACACCGGGCTTTGACAGTGTGACC 2184
 Db 2092 AGACAGAGTGTTCAGGCGCAGGAACCTTCTCCCGGACACCGGAAGTGGCTCTGCACTGTGACCC 2151
 QY 2185 AAGCTGGACTGACACAGCTGTTCTATCGAGATCTGTCTGCCGACTGTGGTGGCCATGG 2244
 Db 2152 TAACTGGACTGCGCCAGACTGCTCAACGAAATATGTTCTGTGGACTGTGGCTCACACGG 2211
 QY 2245 CCGTGTGCTTAGGGGACACTGCGCTGCGAGAGATGGCTGGATGSGGGGAGCTGCGACA 2304
 Db 2212 CTTTGTGATGSGGGGAGAGTGTGCTGTGAAGAGGCTGGACGGGCCCGAGCTGTAAATCA 2271
 QY 2305 GCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGGACCTGCCGCGACGCAAGTGGAGTG 2364
 Db 2272 GAGAGCTGCCACCCCGCTGTGCCGAGCAGCGGACCTGCCGCGACGCAAGTGGAGTG 2331
 QY 2365 CAGCCCTGGCTGGAATGCGGAACACTGACCAATC-----GCTCACTATCTGGATAGGGT 2418
 Db 2332 CAGCCCTGGCTGGAATGCGGAACACTGACCAATCTCCCTAGCTCACTATCTGGATAGGGT 2391
 QY 2419 AGTTAA-----AGAGGGTGGCTTGGCTGGTGTGCAATGGCAACGCGACAGTGTACCTTAGA 2472
 Db 2392 AGTTAACTTTTCAGAGGGTGGCTTGGCTGGTGTGCAATGGCAACGCGACAGTGTACCTTAGA 2451
 QY 2473 CTTGAATGGTTGGCACTGCGCTGCGCTGCGAGTGGGTGGAGAGAGCTGGCTGTGACACTTC 2532
 Db 2452 CTTGAATGGTTGGCACTGCGCTGCGAGCTGGGTGGAGAGAGCTGGCTGTGACACTTC 2511
 QY 2533 CATGAGAGCTGCTGCGCTGACAGCAAGACAATGATGGAGATGGCTGTGGAGTGCAT 2592
 Db 2512 CATGAGAGCTGCTGCGCTGACAGCAAGACAATGATGGAGATGGCTGTGGAGTGCAT 2571
 QY 2593 GGAACCTGACTGCTGCTTCCCTCCAGCCCTGTGCCATATCAACCCGCTGTGCTTGGCTCCCC 2652
 Db 2572 GGAACCTGACTGCTGCTTCCCTCCAGCCCTGTGCCATATCAACCCGCTGTGCTTGGCTCCCC 2631
 QY 2653 TAAACCTCTGAGACATCATCCAGGAGACACAGGTCCTGTGTGCAGAGAAACCTTACATC 2712
 Db 2632 TAAACCTCTGAGACATCATCCAGGAGACACAGGTCCTGTGTGCAGAGAAACCTTACATC 2691
 QY 2713 CTTCTATACCCGCACTCAAGTTCTCTCGTGGCAGGAGACAGCAGCACATAATCCCGGGGA 2772
 Db 2692 CTTCTATACCCGCACTCAAGTTCTCTCGTGGCAGGAGACAGCAGCACATAATCCCGGGGA 2751
 QY 2773 GAAACCCCTTGTATGGAGGCGATGCTGTGTTTATTCGTGGCCAAAGTATGACATCAGATGG 2832
 Db 2752 GAAACCCCTTGTATGGAGGCGATGCTGTGTTTATTCGTGGCCAAAGTATGACATCAGATGG 2811
 QY 2833 AACCCCTGCTGTGTGTAACATCAGTTTGTCAATAACCTCTCTTGGATATCAAT 2892
 Db 2812 AACCCCTGCTGTGTGTAACATCAGTTTGTCAATAACCTCTCTTGGATATCAAT 2871
 QY 2893 CAGCAGCAAGATGGCAGCTTTGACTTGGTGACAAATGGGGGCACTCTCCATCATCTCTGGC 2952
 Db 2872 CAGCAGCAAGATGGCAGCTTTGACTTGGTGACAAATGGGGGCACTCTCCATCATCTCTGGC 2931
 QY 2953 GTTCAGGGGCGACTTTTCATCAGAGGAGCACCCCTGTGGCTGGCATGGGATCGCTT 3012
 Db 2932 GTTCAGGGGCGACTTTTCATCAGAGGAGCACCCCTGTGGCTGGCATGGGATCGCTT 2991
 QY 3013 CTTTGTGATGAAACCATCATCATGAGACATGAGGAGATGAGATTCCAGCTGTGACCT 3072
 Db 2992 CTTTGTGATGAAACCATCATCATGAGACATGAGGAGATGAGATTCCAGCTGTGACCT 3051
 QY 3073 GAGCAATTTTGGCGGCCCCAACCCAGTGTCTCTCATCCCCACTGACGTCTCTCGCCAG 3132
 Db 3052 GAGCAATTTTGGCGGCCCCAACCCAGTGTCTCTCTCATCCCCACTGACGTCTCTCGCCAG 3111

QY 3133 CTCTGTGAGAGAAAGGCCCAATTGTGCGGAAATTCAGGCTTTGAGGAGGAATCTC 3192
 Db 3112 CTCTGTGAGAGAAAGGCCCAATTGTGCGGAAATTCAGGCTTTGAGGAGGAATCTC 3171
 QY 3193 TATCTCTGGTCAAGATGAGGCTGAGCTACCTGAGCAGCGGACCCCTGGCTACAAATC 3252
 Db 3172 TATCTCTGGTCAAGATGAGGCTGAGCTACCTGAGCAGCGGACCCCTGGCTACAAATC 3231
 QY 3253 TGTCTGAGGATCAGCTCACCACCCGACCAATCCCTTTCAACCTCATGAAGGTGCACT 3312
 Db 3232 TGTCTGAGGATCAGCTCACCACCCGACCAATCCCTTTCAACCTCATGAAGGTGCACT 3291
 QY 3313 CATGTTAGCGGTGAGGCGGCTCTCTCAGGAAGTGTTCGTCAGCGCCCAAGACTGTC 3372
 Db 3292 CATGTTAGCGGTGAGGCGGCTCTCTCAGGAAGTGTTCGTCAGCGCCCAAGACTGTC 3351
 QY 3373 CTATTATTTTCATTTGGGACAAAGACAGACGTCTACAAACCAAGAGGTGTTTGGGCTTTCAGA 3432
 Db 3352 CTATTATTTTCATTTGGGACAAAGACAGACGTCTACAAACCAAGAGGTGTTTGGGCTTTCAGA 3411
 QY 3433 AGCTTTGTTTCCGTGGGTATGAATATGAATTCCTCCAGATCTAATCTCTGGGAA 3492
 Db 3412 AGCTTTGTTTCCGTGGGTATGAATATGAATTCCTCCAGATCTAATCTCTGGGAA 3471
 QY 3493 AAGAAACAACAGTCTGAGGCTATGAAATTTGACGCTCCAAAGCTTGGAGGATGAGCT 3552
 Db 3472 AAGAAACAACAGTCTGAGGCTATGAAATTTGACGCTCCAAAGCTTGGAGGATGAGCT 3531
 QY 3553 AGCAAAACATCATGCTCCTCAACATTCAAAGTGTGTTTTCCTGCAAAAGGAAATGGGGA 3609
 Db 3532 AGCAAAACATCATGCTCCTCAACATTCAAAGTGTGCTCCTGCAAAAGGAAATGGGGA 3591
 QY 3610 GATCCAGTTTGTCTCAGCAGCTCTCTGTCATTTGGGAGCATCATGGCAATGGGCGCG 3669
 Db 3592 GATCCAGTTTGTCTCAGCAGCTCTCTGTCATTTGGGAGCATCATGGCAATGGGCGCG 3651
 QY 3670 GAGAAGCATCTCTGCGCCCAAGCTGCAACCGGCTTGTGACGGCAACAAAGCTCTCTGCGCCC 3729
 Db 3652 GAGAAGCATCTCTGCGCCCAAGCTGCAACCGGCTTGTGACGGCAACAAAGCTCTCTGCGCCC 3711
 QY 3730 AGTGGCCCTCAGCTGTGGCTCTGACGGGAGCTCTATGTGGGTGATTTCAATACATTTAG 3789
 Db 3712 AGTGGCCCTCAGCTGTGGCTCTGACGGGAGCTCTATGTGGGTGATTTCAATACATTTAG 3771
 QY 3790 AAGGATCTTCCCTCTGGAATGTCAACCAATCTCAGAGCTGAGG- - - - -AATAAGA 3843
 Db 3772 AAGGATCTTCCCTCTGGAATGTCAACCAATCTCAGAGCTGAGGCTGAGAAATAGA 3831
 QY 3844 TTTCAGACATAGTACAGTCCAGACACAAATACTACCTGGCCACAGACCCCATGAGTGG 3903
 Db 3832 TTTCAGACATAGTACAGTCCAGACACAAATACTACCTGGCCACAGACCCCATGAGTGG 3891
 QY 3904 GGCGCTCTCTTTCTGACAGCAACAGCGCGGGTCTTTAAATCAAGTCCACTGTGGT 3963
 Db 3892 GGCGCTCTCTTTCTGACAGCAACAGCGCGGGTCTTTAAATCAAGTCCACTGTGGT 3951
 QY 3964 GGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTGGGGGACAGGTGACAGGCTCTCC 4023
 Db 3952 GGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTGGGGGACAGGTGACAGGCTCTCC 4011
 QY 4024 CTTTGTAGTACACTGCTGCGGGATGTTGGGAAGGACACAGAGGACCACTCACCAATCC 4083
 Db 4012 CTTTGTAGTACACTGCTGCGGGATGTTGGGAAGGACACAGAGGACCACTCACCAATCC 4071
 QY 4084 CAGGGT- - - - -ATTACAGTGGCAAGTTTGGGCTGATCTACTCTCGTGGATGCGAC 4134
 Db 4072 CAGGGTCCCCAGGCAATACAGTGGCAAGTTTGGGCTGATCTACTCTCGTGGATGCGAC 4131
 QY 4135 CATGATCAGAGGATTCATGAGATGGATCATCTCCAGCTCTGCTGCTCTAATGATCT 4194
 Db 4132 CATGATCAGAGGATTCATGAGATGGATCATCTCCAGCTCTGCTGCTCTAATGATCT 4191

QY 4195 CACATCAGCCCGCCACTCAGCTGTGATTTCTGTATGGATATTTCCAGGTAAAGA- - - - - 4249
 Db 4192 CACATCAGCCCGCCACTCAGCTGTGATTTCTGTATGGATATTTCCAGGTAAAGCAGGT 4251
 QY 4250 - - - - -CTGGAAGTGGCCCAAGACTTACCAATGGAACACTCACTTATGTCT 4305
 Db 4252 TCACCTGGAGTGGCCCAAGACTTACCAATGGAACACTCACTTATGTCT 4311
 QY 4306 CGAACAAATGTGCTGCTGCAAAATCTGAAAAACCAAGGTGCGATTTGTGCGCGGAG 4365
 Db 4312 CGAACAAATGTGCTGCTGCAAAATCTGAAAAACCAAGGTGCGATTTGTGCGCGGAG 4371
 QY 4366 GCCCATGCACTGCCAGGTCCCTGCGATTTGACCACTTTCTGCTTAAAGCAAGGTGCCATCCA 4425
 Db 4372 GCCCATGCACTGCCAGGTCCCTGCGATTTGACCACTTTCTGCTTAAAGCAAGGTGCCATCCA 4431
 QY 4426 CGAACCTTGGAGTCCAGCAGCTTTGGCTGTTTTCACAAATGGGCTCTGTATATTGC 4485
 Db 4432 CGAACCTTGGAGTCCAGCAGCTTTGGCTGTTTTCACAAATGGGCTCTGTATATTGC 4491
 QY 4486 TGAGACTGATGAGAAAAAGATCAACCGCATCAGGAGGTCAACACTAGTGAGAGATCTC 4545
 Db 4492 TGAGACTGATGAGAAAAAGATCAACCGCATCAGGAGGTCAACACTAGTGAGAGATCTC 4551
 QY 4546 ACTGCTTGTGGGCCCCCAGTGTGCTGACTGTAATAATGATGCCAACTGTGATTTT 4605
 Db 4552 ACTGCTTGTGGGCCCCCAGTGTGCTGACTGTAATAATGATGCCAACTGTGATTTT 4611
 QY 4606 TTCTGGAGAGCATGTTTATGCTCAAGGATGCAAAAGTTTAAATACCCCATCTTCTTGGCTGT 4665
 Db 4612 TTCTGGAGAGCATGTTTATGCTCAAGGATGCAAAAGTTTAAATACCCCATCTTCTTGGCTGT 4671
 QY 4666 GTGTGCTGATGGGAGCTCTACGTGGCGGACCTTGGGAAACATCCGAATTCGGTTTATCCG 4725
 Db 4672 GTGTGCTGATGGGAGCTCTACGTGGCGGACCTTGGGAAACATCCGAATTCGGTTTATCCG 4731
 QY 4726 GAAGAAACAAGCTTTCTCAACCCAGAACATGATGAGCTGTCTTACCAATTTACCA 4785
 Db 4732 GAAGAAACAAGCTTTCTCAACCCAGAACATGATGAGCTGTCTTACCAATTTACCA 4791
 QY 4786 GGAGCTCTATCTGTTGATACCAAGGACCTGTACACCAAGCTGCTGCCACAGG 4845
 Db 4792 GGAGCTCTATCTGTTGATACCAAGGACCTGTACACCAAGCTGCTGCCACAGG 4851
 QY 4846 AGACTACCTGTACAACTTCACTGAGGAGCGGAGCATCACACTCATCAGACAA 4905
 Db 4852 AGACTACCTGTACAACTTCACTGAGGAGCGGAGCATCACACTCATCAGACAA 4911
 QY 4906 CAATGGCAACATGTTAAATGTCCCGGAGACTCTTACTGGGATGCCCTCTGGCTGGTGGT 4965
 Db 4912 CAATGGCAACATGTTAAATGTCCCGGAGACTCTTACTGGGATGCCCTCTGGCTGGTGGT 4971
 QY 4966 CCCAGATGGCCAGGTGATCTGGGTGACCATGGGACCAACAGTGCATCAAGAGTGTGAC 5025
 Db 4972 CCCAGATGGCCAGGTGATCTGGGTGACCATGGGACCAACAGTGCATCAAGAGTGTGAC 5031
 QY 5026 CACACAGGACACAGTTGGCATGATGACATACCATGGCAATTCGGCCCTCTGGCCCTGAC 5085
 Db 5032 CACACAGGACACAGTTGGCATGATGACATACCATGGCAATTCGGCCCTCTGGCCCTGAC 5091
 QY 5086 CAAAGCAATGAAACCGGATGGACAAATTTTATGATGACAGAGCTTTGGCCCTGAC 5145
 Db 5092 CAAAGCAATGAAACCGGATGGACAAATTTTATGATGACAGAGCTTTGGCCCTGAC 5151
 QY 5146 AAAATGTGACTTCTTACTGCGCAGGTGAGCAGTTTCCGAGTGTATACAGACAGTTTCAGT 5205
 Db 5152 AAAATGTGACTTCTTACTGCGCAGGTGAGCAGTTTCCGAGTGTATACAGACAGTTTCAGT 5211
 QY 5206 GCATGCTCAGGTAGAGACCTCCAGCAAGGATGTGCCATAACCAACCACTGTCTGC 5265
 Db 5212 GCATGCTCAGGTAGAGACCTCCAGCAAGGATGTGCCATAACCAACCACTGTCTGC 5271
 QY 5266 CTCAGGCGCTTCTTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGG 5325

Db 5272 CTGAGGCGCTTCTACACACTGCTGCAAGACCAAGTCCGGACAGCTACTACATCGGGC 5331
Qy 5326 CGATGGCTCTTGGCGGTGCTGCTGGCCAAACGGCATGGAGGTGGCGCTGCAAGTGGCC 5385
Db 5332 CGATGGCTCTTGGCGGTGCTGCTGGCCAAACGGCATGGAGGTGGCGCTGCAAGTGGCC 5391
Qy 5386 CCACTTGTCTGGCTGACACCGTCAACCCACCGTGGCAAGAGGAATGTCAACGCTGCCAT 5445
Db 5392 CCACTTGTCTGGCTGACACCGTCAACCCACCGTGGCAAGAGGAATGTCAACGCTGCCAT 5451
Qy 5446 CGACAACGGCTCAACCTGTGTGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGGCCAGGT 5505
Db 5452 CGACAACGGCTCAACCTGTGTGAGTGGCGCCAGCGCAAGAGCAGGCTCGGGGCCAGGT 5511
Qy 5506 CACTGTCTTGGCGCGCTGCGGGTG-----CACACCGAAATCTCTATCTCT 5556
Db 5512 CACTGTCTTGGCGCGCTGCGGGTGCTCAGGTTCAACCGAAATCTCTATCTCT 5571
Qy 5557 GGACTTTGATTCGGGTAAACAGCAGAGAGATCTATGATGACACCGCAAGTTACACCT 5616
Db 5572 GGACTTTGATTCGGGTAAACAGCAGAGAGATCTATGATGACACCGCAAGTTACACCT 5631
Qy 5617 TCGGATCTGTAGACGACGCGGGGGCGCCAGCTCTGTGTACCCAGCAGCAGGCTGAA 5676
Db 5632 TCGGATCTGTAGACGACGCGGGGGCGCCAGCTCTGTGTACCCAGCAGCAGGCTGAA 5691
Qy 5677 TGGTGTCAACGCTGACATACCTCCCTGCGGTACATTGTGGCATCCAGAGGGCATCAT 5736
Db 5692 TGGTGTCAACGCTGACATACCTCCCTGCGGTACATTGTGGCATCCAGAGGGCATCAT 5751
Qy 5737 GTCTGAAAGAAATGGAATAGACACGCGGGGCCCATCATCCAGGATCTTCGCTAAGG 5796
Db 5752 GTCTGAAAGAAATGGAATAGACACGCGGGGCCCATCATCCAGGATCTTCGCTAAGG 5811
Qy 5797 GAAGATGAGCTACACATACCTAGAG-----AGTCCATGGTCTGCTACT 5844
Db 5812 GAAGATGAGCTACACATACCTAGAGAGGAGGTGCTCAGTCCATGGTCTGCTACT 5871
Qy 5845 ACACGCGAGGAGCAGTATCTTTGAGTTTCAAGAATGACCGCTCTCTTTCTGTGAC 5904
Db 5872 ACACGCGAGGAGCAGTATCTTTGAGTTTCAAGAATGACCGCTCTCTTTCTGTGAC 5931
Qy 5905 GATGCCCAAGCTGGCGGGGAGACACTAGAGACCATCCGCTCAGTGGGTACTACAGAAA 5964
Db 5932 GATGCCCAAGCTGGCGGGGAGACACTAGAGACCATCCGCTCAGTGGGTACTACAGAAA 5991
Qy 5965 CATCTATCAGCCCTGAGGGCAATGCTCAGTACACAGGACTTCACCTGAGGATGGCA 6024
Db 5992 CATCTATCAGCCCTGAGGGCAATGCTCAGTACACAGGACTTCACCTGAGGATGGCA 6051
Qy 6025 CTTCTTTACACCTTTACTGCGCACTGGCGCAGGGTGATATACAAATATGGCAACT 6084
Db 6052 CTTCTTTACACCTTTACTGCGCACTGGCGCAGGGTGATATACAAATATGGCAACT 6111
Qy 6085 GTCANAAGCTGGCAGAGAGCTCTATGACACCCACCAAGTTCAGTTTCACTATGACAGAC 6144
Db 6112 GTCANAAGCTGGCAGAGAGCTCTATGACACCCACCAAGTTCAGTTTCACTATGACAGAC 6171
Qy 6145 GGCAGGATGCTGAAGACCATCAACCTACAGATGAGGGCTTCACCTGACCATCCGCTA 6204
Db 6172 GGCAGGATGCTGAAGACCATCAACCTACAGATGAGGGCTTCACCTGACCATCCGCTA 6231
Qy 6205 CCGTCAGATGGCCCTGATGACCGACAGATCTTCGCTTCACTGAGGAAGGATGGT 6264
Db 6232 CCGTCAGATGGCCCTGATGACCGACAGATCTTCGCTTCACTGAGGAAGGATGGT 6291
Qy 6265 CAACGCCGTTTTGACTTACAACTATGACAAAGCTTCCGGGTGACACGATGCGGCTGT 6324
Db 6292 CAACGCCGTTTTGACTTACAACTATGACAAAGCTTCCGGGTGACACGATGCGGCTGT 6351
Qy 6325 GATCAACGAGACCCCACTGCGCCATGATCTCTATGCTATGATGATGCTCAGGCAAGAC 6384

Db 6352 GATCAACGAGACCCCACTGCCCCATGATCTCTATCGCTATGATGATGTGTGAGCAAGAC 6411
Qy 6385 AGACAGTTTGGGAAGTTTGTGTGTCATTATATGACANTTAACAGATCATCAACACAGC 6444
Db 6412 AGACAGTTTGGGAAGTTTGTGTGTCATTATGATGACATTACCAGATCATCAACACAGC 6471
Qy 6445 TGTATGATGACCAACCAAGCAATTTTGTATGATATGGCAGGATGAAGAGTGCAGTATGA 6504
Db 6472 TGTATGATGACCAACCAAGCAATTTTGTATGATATGGCAGGATGAAGAGTGCAGTATGA 6531
Qy 6505 GATCTTCGCTCGCTCATGTATCTGATGACCGTCCAGTATGATATAACATGGGGCGAGTAGT 6564
Db 6532 GATCTTCGCTCGCTCATGTATCTGATGACCGTCCAGTATGATATAACATGGGGCGAGTAGT 6591
Qy 6565 GAAAGAGAGCTGAGGTAGGACCTAGCCCAATACCACTCGCTACTCTCTATGATGATGA 6624
Db 6592 GAAAGAGAGCTGAGGTAGGACCTAGCCCAATACCACTCGCTACTCTCTATGATGATGA 6651
Qy 6625 TGCTGACCGCCAGCTGCGAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTA 6684
Db 6652 TGCTGACCGCCAGCTGCGAGACAGTCTCCATCAATGACAAGCCACTCTGGCGCTACAGCTA 6711
Qy 6685 CGACCTCAATGGGAACCTGACCTTACTGAGCCCTGGGAAACAGTGCAGGGCTCACACACT 6744
Db 6712 CGACCTCAATGGGAACCTGACCTTACTGAGCCCTGGGAAACAGTGCAGGGCTCACACACT 6771
Qy 6745 ACGGTATGACATCCGCAACCGCATCACTCGGCTGGGTGACGTGCAATACAAAGATGGATGA 6804
Db 6772 ACGGTATGACATCCGCAACCGCATCACTCGGCTGGGTGACGTGCAATACAAAGATGGATGA 6831
Qy 6805 GGATGGCTTCTGAGGAGCGGGCGGTGATATCTTTGAGTACAACTCAGTGCCTGCT 6864
Db 6832 GGATGGCTTCTGAGGAGCGGGCGGTGATATCTTTGAGTACAACTCAGTGCCTGCT 6891
Qy 6865 CATCAAGCCCTACAAACCGGCTGCGAGTGCAGTACCGCTACAGTGCCTGCT 6924
Db 6892 CATCAAGCCCTACAAACCGGCTGCGAGTGCAGTACCGCTACAGTGCCTGCT 6951
Qy 6925 CGCGCGCTGCTCCAGCAAGAGCAGCCACAGCCACCACTGCACTTCTTATGAGAGCT 6984
Db 6952 CGCGCGCTGCTCCAGCAAGAGCAGCCACAGCCACCACTGCACTTCTTATGAGAGCT 7011
Qy 6985 GACCAACCCACCAAGGTACCCAGCTGTACAAACCACTCAGCTCTGAGATCACTGCT 7044
Db 7012 GACCAACCCACCAAGGTACCCAGCTGTACAAACCACTCAGCTCTGAGATCACTGCT 7071
Qy 7045 CTACTACGACTTGCAGAGGACACCTCTTTGCTGCTGAGCTGAGCTGAGTGTGATGATTTA 7104
Db 7072 CTACTACGACTTGCAGAGGACACCTCTTTGCTGCTGAGCTGAGCTGAGTGTGATGATTTA 7131
Qy 7105 CATAGCTTGTGACAAACATCGGGAACCTCTTTGCTGCTTTAGTGGAAACAGGTTTGTATGAT 7164
Db 7132 CATAGCTTGTGACAAACATCGGGAACCTCTTTGCTGCTTTAGTGGAAACAGGTTTGTATGAT 7191
Qy 7165 CAACCAATCTGTACACAGCTATGGGAGATCTACATGGATACCAACCCCACTTCA 7224
Db 7192 CAACCAATCTGTACACAGCTATGGGAGATCTACATGGATACCAACCCCACTTCA 7251
Qy 7225 GATCATCATAGGCTACCATGGTGGCCCTCTATGATCCCACTCACCAGCTTGTCCACATGGG 7284
Db 7252 GATCATCATAGGCTACCATGGTGGCCCTCTATGATCCCACTCACCAGCTTGTCCACATGGG 7311
Qy 7285 CCGCGAGATTTATGATGTGCTGGCGGAGCGCTGAGCTAGCCAGACCAAGCTGTGGAA 7344
Db 7312 CCGCGAGATTTATGATGTGCTGGCGGAGCGCTGAGCTAGCCAGACCAAGCTGTGGAA 7371
Qy 7345 GCACCTTAGTAGCAGCAACGTCATGCTTTTAACTCTATATGTTTCAAAAACAACACCC 7404
Db 7372 GCACCTTAGTAGCAGCAACGTCATGCTTTTAACTCTATATGTTTCAAAAACAACACCC 7431
Qy 7405 CATCAGCAACTCCCGAGGACATCAAGTGTCTCATGACAGATGTTTAAAGTGTGCTGCTCAC 7464
Db 7432 CATCAGCAACTCCCGAGGACATCAAGTGTCTCATGACAGATGTTTAAAGTGTGCTGCTCAC 7491

QY	7465	CTTTGGATTCCAGCTACACAACGTGATCCCTGGTTATCCCAACACAGACATGATGCCAT	7524
Db	7492	CTTTGGATTCCAGCTACACAACGTGATCCCTGGTTATCCCAACACAGACATGATGCCAT	7551
QY	7525	GGAAACCTCTCTACGAGCTCATCCACACACAGATGAACACCGCAGAGTGGACAAACAGCA	7583
Db	7552	GGAAACCTCTCTACGAGCTCATCCACACACAGATGAACACCGCAGAGTGGACAAACAGCA	7611
QY	7584	-----AGTCTATCTCCGGGGTACAGTGTGAAGTACAGAAGCA	7620
Db	7612	GGTAATTCCCTGCACAAGGCTGCCAGTCTATCTCTCGGGTACAGTGTGAAGTACAGAAGCA	7671
QY	7621	GCTCAAGGCCTTTGTCACTTTAGAACGGTTTGACCAAGCTCTATGGCTTCCCAATCAACCAG	7680
Db	7672	GCTCAAGGCCTTTGTCACTTTAGAACGGTTTGACCAAGCTCTATGGCTTCCCAATCAACCAG	7731
QY	7681	CTGSCACGCGCTCCAAAGACCAAGAGTTTGATCCAGCGGCTCAGTCTTTTGGCAAGGG	7740
Db	7732	CTGCCACGAGCTCCAAAGACCAAGAGTTTGATCCAGCGGCTCAGTCTTTTGGCAAGGG	7791
QY	7741	GGTCAAGTTTCCCTTTGAAGGATGCCAGTGACCCACAGACATCATCAGTGTGCCCAATGA	7800
Db	7792	GGTCAAGTTTCCCTTTGAAGGATGCCAGTGACCCACAGACATCATCAGTGTGCCCAATGA	7851
QY	7801	GGATGGCGGAAGGGTTCTGCCATCTTTGAACCATGCCACTACTCAGAGAACCTGCACCTT	7860
Db	7852	GGATGGCGGAAGGGTTCTGCCATCTTTGAACCATGCCACTACTCAGAGAACCTGCACCTT	7911
QY	7861	CACCAATTGATGGGTGGATACCCATTACTTTGTGAACACAGGACCTTCAGAAGGTGACCT	7920
Db	7912	CACCAATTGATGGGTGGATACCCATTACTTTGTGAACACAGGACCTTCAGAAGGTGACCT	7971
QY	7921	GGCCATCCTGGGCCCTCAGTGGGGGGCGCGAAACCTTGAGAAATGGGGTCAACGTCACGTG	7980
Db	7972	GGCCATCCTGGGCCCTCAGTGGGGGGCGCGAAACCTTGAGAAATGGGGTCAACGTCACGTG	8031
QY	7981	GTCCCAGATCAACACAGTACTTAATGCGAGACTAGACGCTACACAGACATCCAGGTCCA	8040
Db	8032	GTCCCAGATCAACACAGTACTTAATGCGAGACTAGACGCTACACAGACATCCAGGTCCA	8091
QY	8041	GTACGGGGCACTGTGTTTGAACACACGCTACGGGACAACTTGATGAGGAGAGGCACG	8100
Db	8092	GTACGGGGCACTGTGTTTGAACACACGCTACGGGACAACTTGATGAGGAGAGGCACG	8151
QY	8101	GGTCTCGAGCTGGCCGGCAGAGAGCGGTGGCCAGCGCTGGGCCCGCCAGACAGACAG	8160
Db	8152	GGTCTCGAGCTGGCCCGCAGAGAGCGGTGGCCAGCGCTGGGCCCGCCAGACAGACAG	8211
QY	8161	ACTCGCGGAAGGGGAGGAAGCCCTCGCGGGCTTGACAGAGGGGGGAAGCAGAGTGCT	8220
Db	8212	ACTCGCGGAAGGGGAGGAAGCCCTCGCGGGCTTGACAGAGGGGGGAAGCAGAGTGCT	8271
QY	8221	GAGCACAGGGGGGTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTCCAGCAGTACCC	8280
Db	8272	GAGCACAGGGGGGTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTCCAGCAGTACCC	8331
QY	8281	AGAACTGTACAGACAGCGCCAAACATCTCACTTCATGACAGACAGAGCGAGATGGCCGGAG	8340
Db	8332	AGAACTGTACAGACAGCGCCAAACATCTCACTTCATGACAGACAGAGCGAGATGGCCGGAG	8391
QY	8341	GTGACAGAGAGGAC	8354
Db	8392	GTGACAGAGAGGAC	8405

```

1  APPLICANT: Wolenc, Adam R
2  APPLICANT: Vernet, Corine
3  APPLICANT: Eisen, Andrew J
4  APPLICANT: Liu, Xiaohong
5  APPLICANT: Malvankar, Uriel M
6  APPLICANT: Shinkets, Richard A
7  APPLICANT: Tchernev, Velizar
8  APPLICANT: Spaderna, Steven K
9  APPLICANT: Gorman, Linda
10 APPLICANT: Kekuda, Ramesh
11 APPLICANT: Patturajan, Meera
12 APPLICANT: Gusev, Vladimir Y
13 APPLICANT: Gangolli, Esha A
14 APPLICANT: Guo, Xiaojia S
15 APPLICANT: Shenoy, Suresh G
16 APPLICANT: Pastelli, Luca
17 APPLICANT: Casman, Stacie J
18 APPLICANT: Boldog, Ferenc
19 APPLICANT: Burgess, Catherine E
20 APPLICANT: Edinger, Shlomit R
21 APPLICANT: Ellerman, Karen
22 APPLICANT: Gunther, Ezik
23 APPLICANT: Smithson, Glenda
24 APPLICANT: Millet, Isabelle
25 APPLICANT: MacDougall, John R
26
27 TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
28
29 FILE REFERENCE: 21402-230
30 CURRENT APPLICATION NUMBER: US/10/038,854
31 CURRENT FILING DATE: 2003-01-22
32 PRIOR APPLICATION NUMBER: 60/258,928
33 PRIOR FILING DATE: 2000-12-29
34 PRIOR APPLICATION NUMBER: 60/259,415
35 PRIOR FILING DATE: 2001-01-02
36 PRIOR APPLICATION NUMBER: 60/259,785
37 PRIOR FILING DATE: 2001-01-04
38 PRIOR APPLICATION NUMBER: 60/269,814
39 PRIOR FILING DATE: 2001-02-20
40 PRIOR APPLICATION NUMBER: 60/279,832
41 PRIOR FILING DATE: 2001-03-29
42 PRIOR APPLICATION NUMBER: 60/279,833
43 PRIOR FILING DATE: 2001-03-29
44 PRIOR APPLICATION NUMBER: 60/279,863
45 PRIOR FILING DATE: 2001-03-29
46 PRIOR APPLICATION NUMBER: 60/283,889
47 PRIOR FILING DATE: 2001-04-13
48 PRIOR APPLICATION NUMBER: 60/284,447
49 PRIOR FILING DATE: 2001-04-18
50 PRIOR APPLICATION NUMBER: 60/286,683
51 PRIOR FILING DATE: 2001-04-25
52
53 Remaining Prior Application data removed - See File Wrapper
54 NUMBER OF SEQ ID NOS: 411
55
56 SOFTWARE: PatentIn Ver. 2.1
57
58 SEQ ID NO 37
59 LENGTH: 8645
60 TYPE: DNA
61
62 ORGANISM: Homo sapiens
63
64 US-10-038-854-37

```

RESULT 5
US-10-038-854-37
; Sequence 37, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spyttek, Kimberly A
; APPLICANT: Li, Li

QY 206 AGCGCGTCAAGGACATTTGGCCGCGAGGCGCGAGGAATTTCTGCCGACAGGTGCCAAC 265
Db 331 AACAGAGTGAAGATTTGGTTTACAGAGAAGCAGACGAGTTCACTAGACAGAGACAGAAAT 390
QY 266 TTACCCCTCGGAGGTGGGGCTGGAGAAAGTAACGCCCTCACGGGACCCCTGTACCGG 325
Db 391 TTTACCCTAAGGAGTTAGGAGTTTGGAACACGAGCAATCGAAGAGGAGTGGCATTTGT 450
QY 326 ACAGACATTTGGCTGCCCAATGGGTACTTCCATGGGGCTGGCTGTATGCCGACATG 385
Db 451 GCGGAATGGGCTCCCTCACAGAGTTACTTATCAGTCAGGGTCAGATGCTGATCT 510
QY 386 GAGGCTCACACGCTGTCTCCCTGAGCACCCCGTGTGCTGTGGGGCCGAGCACACCG 445
Db 511 GAAATGAAGCAGTGATGTCCCGAGAGCATGCGCATGAGACTTTTGGGCGAGGGGGTCAAA 570
QY 446 TCAGGCGGAGCTCTCGCTGTGACGCGGGCCAAATTCGAATCTCACTACACCGACACC 505
Db 571 TCAGGCGGAGCTCTCGCTGTCAAGTCGGTCAACTCAGCCCTCACCTGTACAGATACG 630
QY 506 GAGCATGAAACACTGAGACTGATCATCCGGGGCGCTGCAGAACACACGGCGGCTCCGG 565
Db 531 GAGCAGAAACAGTCCGACAGTGAGATGAGCAACCTGCAAGCAATCAAGGCCAGTCT 590
QY 566 ACGCCGCGCGCGCTCTCGAGGCGCACACCCCGAACACGACACCGCGGCTCCAT 625
Db 691 ACCCTGAGCGCTTGC CGCTTCCATAGCAGCACTCTGCACAGCATATCCATCCATC 750
QY 626 AACTCTCTGAACCGGGCACTTACGCCGAGAGCAACCCACCGCGCGCCCGACGGAC 685
Db 751 ACTTCTCTCAACAGAACTCCCTGACCAATAGAGGAAACAGAGTCCGGCCCCCGCGCT 810
QY 686 CACTCGCTCTCGGAGAGCCCTTGGCGGGCGGCCAGGAGCTGCCACGCCACAGGAG 745
Db 811 GCTTTGCCCGCGAGCTGCAAC-----CACACCGAGTCCGTCCAGCTGCAGGAC 861
QY 746 AACTGGCTGTCAACAGCAACATCCCTCGAGACACAGAACCTAGCAAGCAGCCATTC 805
Db 862 AGCTGGGTCTTGGCAGTAATGTACCACCTGGAAGCAGGCAATTT----- 905
QY 806 CTAGGACATTCAGGAGCAACCTCATTTAGATGGAATCTTCGGCGCTCCCGCCATGAT 865
Db 906 ----- 905
QY 866 GGGGCTTACAGTGACGGGCACTTCTCTTCAAGCTGGAGGACCTTCCCGCTCTTCTGC 925
Db 906 -----CCTATTCAAAACAGGAAACAGGTACAAACGGCACTGTTCACT 945
QY 926 ACCACATCACAGGTACCCACTGAGTCCAGCACAGTGTACTCTCTCGCGCCCGACCC 985
Db 946 ACTGCAACCCAGATACAAATGGCAATCTGGCTCTGTTTATTCACCACTCTCGGCCA 1005
QY 986 CTGCCCCGACGACCTTTCGCGCGCGGCTTTTAACTCAAGAGCCCTTCAAGTACTGT 1045
Db 1006 CTACCTAGAAACACCCCTATCAAGAGTGTCTTTTAAATTCAGAAAGTCTTCAAGTACTGT 1065
QY 1046 AACTGGAAGTGGGAGCCCTGAGGCGCATCTGTCATCTCAGCACTCTGTCATCTCTGCTG 1105
Db 1066 AGCTGGAATGCACTGCACTGTGTGCGGTAGGGGTCTCGGTGCTCTGCGCAATCTCTG 1125
QY 1106 GCATCTTTCTGGCCATGCACCTGTTTGGCTTAACTGGCACTGCGAGCGATGGAGGG 1165
Db 1126 TCTTATTTTATAGCATGATCTTTTGGCTTCACTGCGAGTACAGCAGACTGAAAT 1185
QY 1166 CAGATGTATGAGATCAGGAGGACACAGCCAGAGTTGGCTGTGCAACCGGACGTCTCC 1225
Db 1186 GACACATTTTGA-----ATGGAAGAGTGAATTTCTGATACCATGCGCAACAACTGTG 1239
QY 1226 CTATACCCCTCAGGGGCACTGCTTAGAGACCCCTGCAGAGGAAAGCAAGAACACCA 1285
Db 1240 TCATTACCTTCTGGAGACA----- 1258

QY 1286 GAAGGAAAGCCAGTAGTTCTTTCCAGAGGACAGTTTTCATAGATTTCTGGAGAAATTCAT 1345
Db 1259 -ATGGAAATTTAGTGGATTTTACGCAAGAAATTAACACCATAGATTTCCGGAGAACTTCAT 1317
QY 1346 GTGGAAAGCGAGCCTCCAGAGATTCCTCTCTGACATTTCTGAGATCTCAAGTGTTC 1405
Db 1318 ATTGCCCAAGAGCAATTCAGAGATTCCTCCGGGATCTTCTGAGATCAAGCTCTTC 1377
QY 1406 ATAGACCATCTGTGCATCTGAATTCATGCTCTGCGGAAAGGAGCCCTGGTTCGC 1465
Db 1378 ATTGATCAGCACAGTTTCTTAAATTCATATCTCTTTCAGAGGATGCAATTCATTCGA 1437
QY 1466 ATTTATGCGAGAAAGGCTCCCTCTTTCACATACACAGTTTTCATTTTGTGGAGTGTCTG 1525
Db 1438 GTATATGCCGGAAGGCTTACCGCTTCCCATACTCAGTATGACTTCTGAGAGCTCTG 1497
QY 1526 GATGCGAGAGGCTCTTAACCCAGAGGCGCGGAGCCTTAGAGGGACCCCCCGCCAGTCT 1585
Db 1498 GATGCGACAGGCTGATTTGCCAGAGAGCAGCGGAGCCTTGTGAGACGGAGAGACCGGG 1557
QY 1586 CGGGAACTGTGCCCTCCAGGCATGAGACAGGCTTTCATCCAGTATTTTCGATTCAGGA 1645
Db 1558 CGGAGGCGAGATCGTCAAGCTTTCATGAGCGCGCTTATCCAGTACTTGGATTCGGA 1617
QY 1646 ATCTGGCACTTGGCTTTTACAATGACGGAAGAGTGCAGAGTGGTTTCTTCTCTCACC 1705
Db 1618 ATCTGGCATCTGGCTTTTATAATGATGGGAAAAATGACAGAGCAGGTCTCTTTAATACC 1677
QY 1706 ACTGCCATTGAGTCGGTGATAAATGCCCCAGCAACTCTATGCAATGTTGACTGCATC 1765
Db 1678 ATTGTATAGAGTCTGTGGTGAATGTCCCGAAATTCGCCATGGAATGGAGATGCGTT 1737
QY 1766 TCTGGCACTGCCACTGCTTCTCGGTTTCTGGGCCCGGACTGTGGCAGAGCCTCTCTGC 1825
Db 1738 TCTGAACTTGCATTGTTTCCAGGATTTCTGGTCCGGATTTGTTCAAGAGCCGCTGT 1797
QY 1826 CCGCTGCTCTAGCGGAATGGCCATACATGAAAGCAGATCTTCTGTCACAGTGC 1885
Db 1798 CAGTGTATGTAGTGGCAACGGGAGTCTTCAAGGGCCGCTGCTGTGTTTCAGCGGC 1857
QY 1886 TGGAAAGCGGTGAGTGCATGTGCCACCAACAGTGTATCGATGTGGCTGCGAGCAAC 1945
Db 1858 TGGAGGCGACCGAGTGTGATGTCCGACTACCCAGTGTATTGACCCACAGTGTGGGGT 1917
QY 1946 CATGCCACTGTATCACGGGCACTGTGATCTGCAACCTTGGCTACAGGGGAGAGCTGT 2005
Db 1918 COTGGGATTTGTATCATGGGCTCTGTGCTTGCAGCTCAGGATACAAAGGAGAAATTTGT 1977
QY 2006 GAGGAAGTGGACTGTGATGCCACCCACATGTTTCAGSCCGGGTGTCTGCTGAGAGCGAA 2065
Db 1978 GAGGAAGTGTACTGTATAGACCTTGGGTGTTCTAATCATGTGTGTGTATCCACGGGAA 2037
QY 2066 TGCCATTCTTTGTGGATGGGAGGACCAACTGCGAGACCCCGAGGCCACATGCTTA 2125
Db 2038 TGTCACTGCAGTCCAGGATGGGAGGTAGCAATTTGAAATACTGAAGACCATGTGTCCA 2097
QY 2126 GACCACTGTTTACGGCCACGGAACTTCTCTCCGGACACCGGCTTTCAGCTGTGACCCA 2185
Db 2098 GACCACTGCTCCGGCCACGGAACTTCTTCAAGAAAGTGGCTCTCTGCAAGTGTGACCCT 2157
QY 2186 AGCTGCACTGACACGACTGTTTCTATCGAGATCTGTGCTGCCGACTGTGTGGCCATGGC 2245
Db 2158 AACTGACTGGCCACAGACTGTCTCAACCAATATGTTCTGTGGACTGTGGCTCACACGGC 2217
QY 2246 GTGTGCTAGGGGCACTTCTGCTGCGAGATGCTGGATGGGGGAGCTGCGACCGAG 2305
Db 2218 GTTTCATGGGGGAGCGTGTGCTGTGAAGAGGCTTGGACGGGCCCCAGCTGTAAATCAG 2277
QY 2306 CGGGCTGCCACCCCGCTGTGCGGAGCATGGGACCTTCCCGCGAGCGCAAGTGCAGTGC 2365
Db 2278 AGAGCTGCGACCCCGCTGTGCGGAGCACGGGACCTGCAAGGATGGCAAGTGTGAATGC 2337
QY 2366 AGCCCTGCTGGAATGGCGAAACATGACCACTATGCTACTATCTGGATAGGGTACT----- 2421

Db	2338	AGCCAGGCTGGAATGAGAGCACTGCACATATCGCTACATATTTGGATAGATAGTTAA	2397
Qy	2422	-----TAAAGAGGTTGCTGCTGGGTTGTGCAATGGCAACGGCAGATGTACC	2467
Db	2398	GACAAGATAGGATATAAAGAGGTTGTCTCTGTGCAACAGCAATGGAAGATGTACC	2457
Qy	2468	TTAGACCTGAATGTTGGCACTGCTGCGCAGCTGGCTGGAGAGAGCTGGCTGTGAC	2527
Db	2458	CTGGACCAAAATGGCGACATTTGTGTGTGTCAGCCCTGGATGGAGAGAGCGCTGTGAC	2517
Qy	2528	ACTTCCATGAGACTGCTCGGTGACAGCAAAAGACAATATGAGATGGCCCTGGTGGAC	2587
Db	2518	GTAGCCATGGAGACTCTTTGCACAGATAGCAGGACAATGAAGGGGATGCACTCATTTAC	2577
Qy	2588	TGCATGGAACCTGACTGCTGCTCCAGCCCTGTGCAATATCAACCGCTGTGCTTGGC	2647
Db	2578	TGCATGGATCCGATTTGCTGCTTACAGAGTTCTGCTCAGATCAGCCCTATTGTGCGGA	2637
Qy	2648	TCCCTTAACCTCTGGACATCATCCAGGAGACACAGGTCCCTGTGTACAGCAGAACCTA	2707
Db	2638	CTGCCGATCCTCAGGACATCATTAGCCAAAGCCTTCAATCGCCTTCTCAGCAAGCTGCC	2697
Qy	2708	CACCTCTTATGACCGCAATAGTTCTCTGCTGGCAGGACAGCAGCAGCATATATCCCC	2767
Db	2698	AAATCCTTTTATGATCGAATCAGTTTCTCTTATAGGATCTGATAGCACCCTATTATACCT	2757
Qy	2768	GGGAGAAACCTTTGATGAGGGCATGCTGTGTATTCTGTGTTATTCGTGGCCAAAGTATGACATCA	2827
Db	2758	GGAGAACTCTTCAATAGAGCCTTGCTGTCTCATCAGAGCCCAAGTACTGCTGCT	2817
Qy	2828	GATGGAACCCCTGTTGTGTGAACATCAGTTTGTGCAATTAACCCCTCTCTTTGGATAT	2887
Db	2818	GATGGAATCCCATTTATGGAGTAAATGTCTGCTTTTTCATTTACCAGCAATATGGATAT	2877
Qy	2888	ACAAATCAGAGGAAGATGCAGCTTTGACTTGTGTCGACAAATGGCGGCATCTCCATCATC	2947
Db	2878	ACTATTACCCGACGAGCAAGATGTTTGTGCTTGGTGGCAATGGTGGGGCTCTCTAATC	2937
Qy	2948	CTGGGTTGAGCGGCACCTTTTATCACAAGAGACACACCTGTGGTGGCATGGAT	3007
Db	2938	TTGTATTTGAACGATCCCATTTCTCCTCAGTATCATATCTGTGTGGATTCATGGAAT	2997
Qy	3008	CGCTTTTGTGTAAGAACCATCATCAGACATGAGAGATGAGAGATCCAGCTGT	3067
Db	2998	GTCTTTTATGTGATGATACCTTAGTATGGAGAAAGAGAAATGACATTTCCAGCTGT	3057
Qy	3068	GACCTGAGCAATTTTGGCCGCCCCAACCCAGTGTCTCTCCATCCCCCACTGACGTCTTTC	3127
Db	3058	GATCTGATGATTTCTGTGGGCCAAATCCCATATTGTGTATCATCTTTATCCACCTTT	3117
Qy	3128	GCAGCTCTCTGTGAGAGAAAGCCCATTTGTGCGGAAATTCAGGCTTTGAGAGGAA	3187
Db	3118	TTGAGATCTTCTCTGAGAGACATCCCATCATTTCCCGAAACACAGGTACTCCACGAGAA	3177
Qy	3188	ATCTCTATCTGCTGCTGCAAGATGAGGCTGAGTACTCTGAGCAGCGCGACCCCTGGCTAC	3247
Db	3178	ACTACAAATCCAGGACAGATTTGAACTCTCTCTACTTGTGATTCAGAGCTGCAGGATAT	3237
Qy	3248	AAATCTCTCTGAGGATCAGCTCAGCCACCCAGCCATCCCTTCAACCTCATGAGGTT	3307
Db	3238	AAGTCAGTTCTCAAGATCAACATGACCCAGTCTATTATTCCATTTAATTAATGAGGTT	3297
Qy	3308	CACCTCATGTAGGTTGAGGCGCCCTCTTCCAGGAAGTGTTCCTGCAGCCCCAGAC	3367
Db	3298	CATCTTATGTAGTGTAGTAGAAGACTCTTCCAAAGTGTGTTCTGCTCATCACCAC	3357
Qy	3368	CTGTCTTATTTATTTTGGGCAAGACAGCTCTTACAACAGAGAGTGTGTGGGCTT	3427
Db	3358	TTGGCCCTATCTTTCATATGGGATAAAACAGATGTCATATAATCAGAAAGTCTATGTCTA	3417
Qy	3428	TCAGAGCCCTTGTTCCTGGGTTATGATATGATATCTGCTGCCAGATCTAATCTGCTGG	3487

Db	3418	TCTGAAGCTGTTGTGTGAGTTGGATATGAGTATGAGTCGTGTTTGGACCTGACTCTGTGG	3477
Qy	3488	GAAAAAGACAACAGTGTGCGGGCTATGAATTTGACCGCTCCAAAGCTTTGGAGATGG	3547
Db	3478	GAAAAGAGGATGTCATTTCTGCAAGGCTATGAATTTGGATCGCTCCAAACATGGGTGGCTGG	3537
Qy	3548	AGCCTAGACAAAACATCATATGCCCTCAACATTTCAAAGTGTATCTCTGCACAAAGGAATGGG	3607
Db	3538	ACATTTAGATAAACAATCAAGTGTGTACAGAACGTTATCTGTACAAGGAAACGGG	3597
Qy	3608	GAGAACAGATTTGTGTCTCAGCAGCTCTCTGTCTATTGGGAGCATCATGGGCAATGGCGC	3667
Db	3598	GAAAACAGATTTCACTCTCCAGCAGCTCCAGTCTGTAGTAGCATATGGGCAATGGGCGA	3657
Qy	3668	CGGAGAGCATCTCTCTGCCCCAGCTGCAACGGCTTCTCTGACGCGCAACAAAGCTCTTGCC	3727
Db	3658	AGGCGAGCATTTCTCTGCCCATTTGCAATGGTCAAGCTGATGGTAACAAGTTACTTGCC	3717
Qy	3728	CCAGTGGCCCTCACTGTGGCTCTGACGGAGCCCTCTATGTGGGTATTTCAATACATTT	3787
Db	3718	CCAGTGGCGCTAGCTTTGTGGGATCGATGGCAGTCTGTACGTAGGCGGATTTCAATACGCTG	3777
Qy	3788	AGAAAGATCTTCCCTCTTGGAAATGTCCCAACATCTAGAGCTGAGGAATAAAGATTTC	3847
Db	3778	CGGCGGATATTTCCCTCTTGGAAATGTAAACAAGTCTTTAGAACTAAGAAATAAAGATTTT	3837
Qy	3848	AGACATAGTCAAGTCCAGCACACAAATACTACTGGCCACAGACCCCATAGTGGGGCC	3907
Db	3838	AGACATAGCAGCAACCCAGCTCATAGATCTACTCTTCAACCGGATCCAGTCACGGAGAT	3897
Qy	3908	GTCTTCTTCTTGACAGCAACAGCCGGGCTCTTTAAATCAAGTCCACTGTGTGGTG	3967
Db	3898	CTGTACGTTTCTGACACAAACACCCGAGAAATTTATCGCCCAAGTCTACTTACGGGGCA	3957
Qy	3968	AAGACCTTTGTCAAGAACTCTGAGTGGTTGCGGGGACAGGTGACCACTGCTCCCTTT	4027
Db	3958	AAAGCTTGACTAAATATGCAAGATCTGTCGAGGACAGGGAGCAATGCCCTCCGTTT	4017
Qy	4028	GATGACACTGCTCGGGGATGGTGGAAAGCCACAGAAAGCCACACTCACCATATCCAGG	4087
Db	4018	GACGAGCGAGATGTGGGATGGAGGAAGCCGCTGGAAGCCACACTCATGAGTCCCAA	4077
Qy	4088	GGTATTACAGTGGACAAAGTTTGGCTCATCTACTTCTGTGATGGCAACCATGATCAGACGC	4147
Db	4078	GGAATGCGAGTTGTATGAATGGATTAATCTACTTTGTTGATGGAACCATGATTAGAAA	4137
Qy	4148	ATCATCAGAAATGGATCATCTCCACCTGCTCGGCTCTAATGATCTCATCAGCCCGG	4207
Db	4138	GTTCACAAAATGGAATCATATCAACTCTTCTGGGCTCTAACGATTTGACTTCAGCCAGA	4197
Qy	4208	CCACTCAGCTGTGATTTCTGTCATGATATTTCCAGGTAAAGCTGAGTGGCCACAGAC	4267
Db	4198	CCTTTAATTGTGACACAGCATGCAATCAGCCAGGTACGTCTGGAATGGCCACCTGAC	4257
Qy	4268	TTAGCCATCAACCAATGGAACACTCTATTATGTCTCCAGCAACAAATGTGGTCTGCAA	4327
Db	4258	CTAGCCATTAACCTTATGGATTAATCCATTTATGTCTCGGATAATAATGATGTTTACAG	4317
Qy	4328	ATCTCTGAACACACAGGTGCGCATTTGTGCGGGAGGCGCCATGCTACCTCCAGGTCCCT	4387
Db	4318	ATCATGAAATCTGTAAGTTCGATTTGCTGTGGAGGCGCCATGCTACCTGAGGTTCCC	4377
Qy	4388	GGCATTGACCACTTCTCTGCTAAGCAAGGTGGCCATCCACGCAACCCCTGGAGTCAGCCACC	4447
Db	4378	GGAGTGG---AATATCTGTGGGAAGCACGCGGTGCAGACAACTGGAATCAGCCACT	4434
Qy	4448	GCTTTGGCTGTTTCAACAATGGGCTCTGTATATTCTGAGACTGATGAGAAAAGATC	4507
Db	4435	GCCATTGTGTGCTCTCAGTGGGGTCTCTGTACTATTACTGAAACTGATGAGAAATAAT	4494
Qy	4508	AACCGCATCAGGAGGTCAACACTTAGTGGAGATCTCACTCTGCTGGGCGCCCGCCAGT	4567
Db	4495	AACCGGATAGGCGAGGTCAACAACAGATGGAGAAATCTCTTAGTGGCGGAAATACCTTCA	4554

QY 4568 GGCTGTGACTGTAAAAATGATGCCAACTGTGATGTTTTTCTTGAGACGATGTTATGCC 4627
 Db 4555 GAGTGTGACTGCAAAAAATGATGCCAACTGTGATGTTTTTACCCAGAGTGAGATGCTACGCC 4614
 QY 4628 AAGGATCCAAAGTTAAATAACCCATCTTCTGCTGTGTCTGCTGATGGGAGCTCTAC 4687
 Db 4615 AAGATGCCAAACTCAGTGGCCCATCTCTCTGCTGTGTCTCTCAGATGTAACATGTAT 4674
 QY 4688 GTGGCCGACCTTGGGAACATCCGAATTCGGTTTTATCCGGAAGAACAGCCCTTCCCTCAAC 4747
 Db 4675 ATTGCAGATCTAGGGAATATCCGGATCCGGCTGTGTCAAAAGAAATAGCCCTTTACTTAAC 4734
 QY 4748 ACCCAGAACATGATGAGTGTCTTACCAAAATGACAGGAGCTCTATCTGTTTGTATACC 4807
 Db 4735 TCTATGAATCTTATGAAGTTGGCTCTCCAACTGATCAAGAACTCTACATCTTTTGACATC 4794
 QY 4808 ACCGGCAAGCACCTGTACACCAAGCCTGCCACAGAGACTACCTGTACAACTTCACC 4867
 Db 4795 AATGGTACTCACCAATATATCTGAAGTTTGTAGTCACTGGTGATTAACCTTTTACATTTAGC 4854
 QY 4868 TACACTGGGAGCGCGCATCACACTCATCACAGAACAAATGGCAACATGSGTAAATGTC 4927
 Db 4855 TACAGCAATGACAATGATATTAATCTGTGTGACAGACAGCAATGCCAACACCCCTTAGAAT 4914
 QY 4928 CGCGAGACTCTACTGGGATGCCCTCTGGCTGTGTGCTCCAGATGCCAGGTGTACTGG 4987
 Db 4915 AGACGGACCAATCCATGCCAGTTTCGAGTGTGTCTCTGATACCAAGTGAATGG 4974
 QY 4988 GTGACCATGGGCACCAACAGTGCACTCAAGAGTGTGACCAACACAAAGACACGAGTTGGCC 5047
 Db 4975 TTGACAATAGGAACAAATGGATGTTTGAAGGCATGACTGCTCAAGGACTGGAATTAGTT 5034
 QY 5048 ATGATGATACCAATGCAATTCGGGCTTCTGGCAACCAAAAGCAATGAACGATGAG 5107
 Db 5035 TTGTTTACTTACCAATGCAATGAGTGGCTTTTACCACTAAAGTGAATGAATGAGTGG 5094
 QY 5108 ACAACATTTTATGATGACGAGCTTTGGCGGCTGACAAATGTGACCTTCCCTACTGGC 5167
 Db 5095 ACAACGTTTTTGACTATGACAGTGAAGTGTCTGACAAATGTACGTTTCCACTGGA 5154
 QY 5168 CAGTGAGCAGTTTCCGAATGATACAGACAGTTTCAGTGTGATTCAGGTGAGAGACCTCC 5227
 Db 5155 GTGGTCAAAACCTGATCGGAGCATGGACAAGGCTATCACAGTGGACATGAGTCAATCT 5214
 QY 5228 AGC---AAGGATGATGTCACATAACCAACCACTGTCTGCTCAGCGGCTTCTACACA 5284
 Db 5215 AGCCGAGAAGAGATGTCAGATCACTTCAAATCTGTCTCGATCGATCTTTCTTACACC 5274
 QY 5285 CTGCTGCAAGACCAAGTCGGAAACAGCTACTACATCGGGCCGATGCTCTTTCGGGCTG 5344
 Db 5275 ATGGTTCAAGATCAGTTAAGAAACAGCTACCAAGTTGTTATGACGGCTCCCTCAGAAAT 5334
 QY 5345 CTGCTGCGCAACGGCATGGAGTGGGCTCGAGACTGAGCCCACTTGTCTGGCTGGCCACC 5404
 Db 5335 ATCTACGCCAGTGGCTTGGACTCACACTCAACACAGAGCCGACCGTTCTGGCTGGCAC 5394
 QY 5405 GTCAACCCCAACCGTGGCAAGAGAAATGTCAACGCTGCCATCGACAAACGGCTTCAACCTG 5464
 Db 5395 GCTAATCCGACGGTTGCCAAAGAAACATGACTTTGCCCTGGCGAAGAGGTCAAACCTG 5454
 QY 5465 GTGAGTGGCGCCAGCCGCAAGAGCAGCTCGGGGCCAGGTCTACTGTCTTGGCGCCGG 5524
 Db 5455 GTGGAATGGAGATTCGAAAGAGCAAGCCCAAGGAAAGTCAATGTCTTTGGCCGCAAG 5514
 QY 5525 CTGCGGTGCACAACCGAAATCTCTATCTCTGACTTTGATCGCGTAAACACCAAGCAGAG 5584
 Db 5515 CTGAGGTATGAGCAGAAACCTTCTTTCAGTTTGAATTTGATCGAAACCAAGACAGAA 5574
 QY 5585 AAGATCTATGATGACCAACGAAAGTTCACCTTTCCGATTTCTGACCAAGCGGGCGG 5644
 Db 5575 AAGATCTATGACGACCAACCGTAAATTTCTACTGAGGATCGCTTACGACAGCTCTGGGCAC 5634

QY 5645 CCCACGCTCTGTGTACCCAGCAGCAGGCTGAATGTGTCAACGTGACATACTCCCTGGG 5704
 Db 5635 CCGACTCTCTGGCTGCCAAGCAGCAAGCTGATGGCCGTCAATGTCACTTATTCATCCACA 5694
 QY 5705 GGTTCACATTTGGTCATCCAGGGGGCATCATGTCTGAAGAAATGGAATACGACACGCG 5764
 Db 5695 GGTCAAATTCGAGCATCCAGGAGCACCACTAGCGAGAAAGTAGATATGACGGAACAG 5754
 QY 5765 GGGCCGATCACAATCCAGGATCTTCGCTGATGGGAAGACATGAGCTACACATCTTAGAG 5824
 Db 5755 GGGAGGATCGTGTCTCGGGTCTTCTGCTGATGGTAAACATGAGTTACACATATTTAGAA 5814
 QY 5825 AAGTCCATGGTCTGCTACTACAGCCAGGAGGAGTATATCTTTGAGTTCGACAGAAT 5884
 Db 5815 AAGTCCATGGTCTTCTGCTTATAGCCAGCGGAGTACATCTTCGAATACGATATGTGG 5874
 QY 5885 GACCCGCTCTCTTCTGTGACGATGCCCAACGTGGCGGCGACACACTAGAGACCATCCGC 5944
 Db 5875 GACCGCTCTGCTGCCATCACCATGCCAGTGTGGCTGCCACACATCGACGACCATCCGA 5934
 QY 5945 TCAGTGGGCTACTACAGAAACATCTATCAGCCCTTGAGGGCAATGCTCTAGTCTATACAG 6004
 Db 5935 TCCATTGGCTACTACCGCAACATATACAAACCCCGGAAAGCAACGCTTCCATCATCAG 5994
 QY 6005 GACTTCACTGAGGATGGGCACCTCTTCCACCTTCTACAGAGTCTTCTGGGTACAAGTCGGAGGTC 6064
 Db 5995 GACTACAAGAGAGAGGCTGCTTCTACAACAGTCTTCTGGGTACAAGTCGGAGGTC 6054
 QY 6065 ATATACAAGTATGGCAACTGTCAAGCTGCGAGAGCGCTCTATGACACCAACCAAGTCT 6124
 Db 6055 TTATTTCAAATACAGAAAGCAGACTAGGCTCTCAGAAATTTTATATGATAGACAAAGTCT 6114
 QY 6125 AGTTTCACTATGACAGACGCGGCGCATGTGAGACCATCAACCTACAGATGAGGCG 6184
 Db 6115 AGTTTTCCTATGATGAACAGCAGGAGTCTTAAGACAGTAAACCTCCAGAGTATGGT 6174
 QY 6185 TTCACTCGCACCATCCGCTACCGTACAGATTTGGGCGCTGATTTGACCGCAGATCTTCGC 6244
 Db 6175 TTATTTGCACCATTAGATACAGGCAATTTGGTCCCTGATTTGACAGCAGATTTTCGC 6234
 QY 6245 TTCCTAGGAAAGGATGGTCAACGCCCTTTTGAATCAACATATGACAAAGCTTCGG 6304
 Db 6235 TTAGTGAAGATGGGATGGTAAATGCAAGATTTGACTATAGCTATGACACAGCTTTCGA 6294
 QY 6305 GTGACGACATGACGAGCTGTGATCAACGAGACCCCACTGCCATTTGATCTCTATCGCTAT 6364
 Db 6295 GTGACGACATGACGAGGCTGTGATCAATGAACGCACTGCCCTATTTGATCTGTATCAGTT 6354
 QY 6365 GATGATGTGTGAGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCAATTTACTATGACAT 6424
 Db 6355 GATGACATTTCTGGCAAGTTGAGCAGTTTGGAAAGTTTGGAGTTATATATTTATGATAT 6414
 QY 6425 RACCAATCATCACACAGCTGTGATGACCCACACCAAGCATTTTGTATGATATGCGCAGG 6484
 Db 6415 RACCAATCATTTTACAGCTGTATGACCTTATGACCTTATGACAGCAGCTTTGATGCTCATGGCCGT 6474
 QY 6485 ATGAAGGAAGTGCAGTATGAGATCTTCGCTCGCTCATGTACTGATGACCGTCCAGTAT 6544
 Db 6475 ATCAAGGAGATTTCAATATGAGATTTCAAGTGTCTCATGTACTGATGATTAATTTCAAT 6534
 QY 6545 GATACATGGGCGAGTGTGAGAGGAGCTGAAGGTAGGACCTTACGCAATTCCTACT 6604
 Db 6535 GATAATAGGCTGGGTAAACCAAGAGAGAGATTTAAATAGGGCCCTTTTGGCAACCAACC 6594
 QY 6605 CGCTACTCTGTAGTGTGATGCTGACGGCAGCTGACAGAGCTCTCCATCAATGACAAG 6664
 Db 6595 AAATATGTTTATGATATGATGTTGATGGAAGCTTCCAAACAGTTTACTCAATGAAAAG 6654
 QY 6665 CCACCTCTGGGCTACAGCTACGACTCAATGGGAACCTGCACTTACTGAGCCCTGGGAAC 6724
 Db 6655 ATAAATGGCGGTACAACTAGCATCTGAATGGAACCTTCCATTTTACTGAACCAAGTAAAC 6714
 QY 6725 AGTGCAGGCTCACACCACTACTAGGTATGACATCCGCGACCGCATCACTCGGCTGGGTGAC 6784

Db 6715 AGTGGCGTCTGACACCCCTTCGATGACCTCGAGACAGAAATCACTCGACTGGGTGAT 6774
 Qy 6785 GTGCAATACAAAGATGAGTGGCTTCTGAGGAGCGGGCGGTGATATCTTTGAG 6844
 Db 6775 GTTCAATATCGTTGGATGAGATGGTTCTAGCTCAAGGGGACGGAATCTTTGAA 6834
 Qy 6845 TACAACCTCAGCTGGCTCTCATCAAGGCGCTACAACCGGCTCGAGCTGGAGTGTGAG 6904
 Db 6835 TATAGCTCAAGGGGCTTCTAACTCGAGTTTACAGTAAAGGCGAGTGGCTGGACAGTATC 6894
 Qy 6905 TACCGCTACAGTGGCTGGGCGGCGGTGCTCAGCAAGAGCGACACAGCCACCTG 6964
 Db 6895 TACCGTTATGACGGCTTGGGAAGCGGTGTTTCTAGCAAAACCAAGTCTAGGACAGCACCTG 6954
 Qy 6965 CAGTTCTTCTATGACAGCTGACCAACCCCAACCAAGTCAACCACTGTACAACTCC 7024
 Db 6955 CAGTTTTTTATGCTGACTTAATCTTATCCACTAGGATTACTCATGTCTACAACTTCG 7014
 Qy 7025 AGCTCGAGATCACTCCCTCTACTAGACTTGCAGAGACACCTCTTTGGCATGGAGCTG 7084
 Db 7015 AGTTTCAAAAATACCTCCCTCTATTATGATCTCAAGGACATCTTTTGGCATGGAAATC 7074
 Qy 7085 AGCAGTGGTATGATGTTTACATAGCTGTGACAAATCGGACCCCTCTTGTGCTTT 7144
 Db 7075 AGCAGTGGGATGATTTCTATATTGCTATCGATTAACAGGGACACCATGCTGTGTTTC 7134
 Qy 7145 AGTGAACAGGTTTGATGATCAAGCAAACTCTGTACACAGCTTATGGGAGATCTACATG 7204
 Db 7135 AGTAGCAATGGCTTATGCTGAAACAGATTCACTGATGATGATGGGAAATCTATTTT 7194
 Qy 7205 GATACCAACCCCACTTTCAGATCACTAGGCTACCTGCTGGCTCTCTATGATCCACTC 7264
 Db 7195 GACTCTAATATTGACTTTCAGTGTGTAATGGATTTTCATGGTGGCTGTATGACCCACTC 7254
 Qy 7265 ACCAAGCTTGTCCACATGGGCGCGGAGATTATGATGTGCTGGCGGAGCGCTGACTAGC 7324
 Db 7255 ACCAAATTAATCCACTTTGAGAAAGAGATTATGACATTTTGCAGGCGGTGACAA 7314
 Qy 7325 CCAGACACAGCTGTGGAGACCTTAGTAGAGCAACGTCTATGCTTTTATCTCTAT 7384
 Db 7315 CTTGACATAGAAATCTGGA---AAGAAATTTGGAAGGACCCAGCTCCTTTTAACTTGTAC 7371
 Qy 7385 ATGTTTCAAAACCAACCCCATCAGCAACTCCAGGACATCAAGTCTTCAATGACAGAT 7444
 Db 7372 ATGTTTAGGAATAACAACCTCGAAGCAAAATCCATGACGTGAAGATTACATCAGAT 7431
 Qy 7445 GTTAACAGCTGGTGTCTCACTTTGGATTCAGCTACAAAGCTGATCCCTGGTTATCC 7504
 Db 7432 GTTAACAGCTGGTGTGATGATTTGGTTTCCATCTGCACAAATCTTCTGCTGATCCCT 7491
 Qy 7505 AAACCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7564
 Db 7492 GTTCCCAATTTGATTTAAACAGAACCTTTTACGA-----ACTTGTGAAGAT 7539
 Qy 7565 CAGAGTGGGACAAACAGAACTTATCTCTCGGGGTACAGTGTGAAGTACAGAAAGCTC 7624
 Db 7540 CAGCAGTGGGATGATATACCGCCATCTTCGGAGTCCAGCAGCAAGTGGCGGCGAGGC 7599
 Qy 7625 AAGCCCTTTGACCTTAGAACCGTTGACAGCTCTATGGCTCCACATCACCAGCTGC 7684
 Db 7600 AAGCCCTTCTGTGCTGGGGAAGATGGCGGAGTGGAGGTGAGCTGAGCGCGCGGCGGCG 7659
 Qy 7685 CAGCAGCTCCAAAGACCAAGAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGCTC 7744
 Db 7660 ---GGCGCGAGTCTCGCTGTGTTTCGCCACCGTCAAGTGGCTGATCGCAAGGCGCTC 7716
 Qy 7745 AAGTTTCCCTTGAAGGATGGCCAGTGAACCAAGATCATCATGTTGGCCATGAGGAT 7804
 Db 7717 ATGCTGGCCCTGACGCCAGGCGCGTGCAGACCAACGTGCTCAACATCGCAACAGGAC 7776
 Qy 7805 GGGCGAAGGGTGTGCTGCATCTTGAACCAATGCCACTTACCTAGAGAACTGCACTTCCAC 7864

Db 7777 TGCAATCAAGTGGGCGGCGGTGCTCAACAACGCTTCTACCTGGAGAACCTGCACTTCAAC 7836
 Qy 7865 ATTGATGGGTGGATATCCATTACTTTTGTGAACACGAGACCTTCAGAAAGTGTACCTGGCC 7924
 Db 7837 ATCAGGGCAAGGACACGCACTATTTCATCAAGACACACCGCCGAGAGACCTGGCC 7896
 Qy 7925 ATCTGGGCTCAGTGGGGGCGGCGAAACCTTGAGAAATGGGGTCAACGTCACTGTGTCC 7984
 Db 7897 ACGTGGGTGTTGACAGCGGCGCAAGCGCTGAGAAACGCGCATCAACGTGACGGTGTG 7956
 Qy 7985 CAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCAGTAC 8044
 Db 7957 CAGTCCACCAGGTGTGAACCGGACGACGCGCAGGTTTCGGGACGTGGAGATGCACTTC 8016
 Qy 8045 GGGCACTGTGCTTTGAAACACACGCTACGGGACAAACGTTGGATGAGGAGAAAGGCAAGGTC 8104
 Db 8017 GGGCGCTGGCGCTGCACGTGCGCTACGCGATGACCTGGACGAGGAGAAAGGCGGCATC 8076
 Qy 8105 CTGAGCTGGCGCGGCGGACAGAGCGCTGCGCAACGCTGGCGCGGCGGCGGACGAGACTG 8164
 Db 8077 CTGAGCAGGCGCGGCGGAGCGCGCTCGCCGGGCTTGGCGCGGCGGAGCAGCGGCTG 8136
 Qy 8165 CGGAAAGGGGAGGAGGCGCTCGCGGCGCTGGAACAGAGGGGAGAGACGAGGTGCTGAGC 8224
 Db 8137 CGCAGCGCGGAGGAGGCGCGGCGCTCTGGAAGGAGGCGGAGAGCGGCGAGCTGCTGAGC 8196
 Qy 8225 ACAGCGGGTGTGAAGGCTAGCAGGCTTTTCTGTGATCTCTCTGAGAGCTACCCAGAA 8284
 Db 8197 GCGGCAAGGTGCGAGGCTAGCAGGCTTCTGCTGAGAGCTACCTGCTGAGAGCTACCCGAG 8256
 Qy 8285 CTGTCAGACGCGCAACACATCCACTTATGAGACAGAGGAGAGTGGCGCGGAGGTGA 8344
 Db 8257 CTGCGGACGCGGCAACACATCCAGTTCTGCGGAGAGGAGATCGGAGGAGTAA 8316
 Qy 8345 C 8345
 Db 8317 C 8317

RESULT 6

US-10-038-854-35
 ; Sequence 35, Application US/10038854
 ; Publication No. US20040022781A1

GENERAL INFORMATION:

; APPLICANT: Spyttek, Kimberly A
 ; APPLICANT: Li, Li
 ; APPLICANT: Wolenc, Adam R
 ; APPLICANT: Vernet, Corine
 ; APPLICANT: Eisen, Andrew J
 ; APPLICANT: Liu, Xiaohong
 ; APPLICANT: Malyankar, Uriel M
 ; APPLICANT: Shimkets, Richard A
 ; APPLICANT: Tchernev, Velizar
 ; APPLICANT: Spaderna, Steven K
 ; APPLICANT: Gorman, Linda
 ; APPLICANT: Kekuda, Ramesh
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Gusev, Vladimir Y
 ; APPLICANT: Gangolli, Esha A
 ; APPLICANT: Guo, Xiaojia S
 ; APPLICANT: Shenoy, Suresh G
 ; APPLICANT: Rastelli, Luca
 ; APPLICANT: Casman, Stacie J
 ; APPLICANT: Boldog, Ferenc
 ; APPLICANT: Burgess, Catherine E
 ; APPLICANT: Edinger, Shlomit R
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Smithson, Glenda
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: MacDougall, John R
 ; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
 ; FILE REFERENCE: 21402-230

	;	CURRENT APPLICATION NUMBER:	US/10/038,854	
	;	CURRENT FILING DATE:	2003-01-22	
	;	PRIOR APPLICATION NUMBER:	60/258,928	
	;	PRIOR FILING DATE:	2000-12-29	
	;	PRIOR APPLICATION NUMBER:	60/259,415	
	;	PRIOR FILING DATE:	2001-01-02	
	;	PRIOR APPLICATION NUMBER:	60/259,785	
	;	PRIOR FILING DATE:	2001-01-04	
	;	PRIOR APPLICATION NUMBER:	60/269,814	
	;	PRIOR FILING DATE:	2001-02-20	
	;	PRIOR APPLICATION NUMBER:	60/279,832	
	;	PRIOR FILING DATE:	2001-03-29	
	;	PRIOR APPLICATION NUMBER:	60/279,833	
	;	PRIOR FILING DATE:	2001-03-29	
	;	PRIOR APPLICATION NUMBER:	60/279,863	
	;	PRIOR FILING DATE:	2001-03-29	
	;	PRIOR APPLICATION NUMBER:	60/283,889	
	;	PRIOR FILING DATE:	2001-04-13	
	;	PRIOR APPLICATION NUMBER:	60/284,447	
	;	PRIOR FILING DATE:	2001-04-18	
	;	PRIOR APPLICATION NUMBER:	60/286,683	
	;	PRIOR FILING DATE:	2001-04-25	
	;	Remaining Prior Application data removed - See File Wrapper or PALM.		
	;	NUMBER OF SEQ ID NOS:	411	
	;	SOFTWARE:	PatentIn Ver. 2.1	
	;	SEQ ID NO 35		
	;	LENGTH:	8675	
	;	TYPE:	DNA	
	;	ORGANISM:	Homo sapiens	
	;	US-10-038-854-35		
		Query Match	41.6%; Score 3473.8; DB 17; Length 8675;	
		Best Local Similarity	65.2%; Pred. No. 0;	
		Matches 5445; Conservative	0; Mismatches 2692; Indels 216; Gaps 14;	
Qy	35	ATGAGCGTGAAGAGAGAGACCTTACCCTGCTGACCCGGCC---CGGACCCGAG	91	
Db	151	ATGGATGTGAAGAACCGCAGGCCTTACTGCTCCCTGACCAAGCAGACGAGAAGAA	210	
Qy	92	CGCGCTACACCAAGCTGTCGCGGACAGCGAGGAGGCAAAG---CCCGCAGAAAATCG	148	
Db	211	CGSGGCTACACAATTTCTTCCGAGACATGAGGAGTGCGGGTACCACACAGAGTCC	270	
Qy	149	TACAGTCCAGCGAACCCTGAAGGCTACGACCAGGA---CGCCGCCCTAGCCATTGGC	205	
Db	271	TACAGTTCACGCGACATTTGAAGCTTTTGTATCATGATTCCTCGCGGCTGCTTTACGGC	330	
Qy	206	AGCGCGTCAAGACATTTGCGCGAGGAGCGCGAGGATTTCTGCGCACAGTCCCAAC	265	
Db	331	AACAGAGTAGAGATTTGGTTACAGAGAAGCAGACGAGTTCACTAGACAGGACAGAT	390	
Qy	266	TTCAACCTCGGAGCTGGGGCTGGAAAGATAACGCCGCCCTCACGGACCCCTGTACGG	325	
Db	391	TTTTACCTTAAGCGAGTTAGAGTTTGTGAACAGCAACTCGAAGAGGACTGGCAATTTGT	450	
Qy	326	ACAGACATTTGGCTGCCCAATCGGCTACTCATGGGGCTGGCTCTGATGCCACATG	385	
Db	451	GCGGAAATGGGCTCCCTCACAGAGTTACTTATCAGTGCAGGGTACAGATGCTGATACT	510	
Qy	386	GAGGCTGACACGCTGTGCTCCCTGAGCAACCCCGTGCCTGTGGGGCGGAGCACAGG	445	
Db	511	GAATAAGACAGTATGTTCCCAGAGCATGCCATGAGACTTTGGGCGAGGGGGTCAA	570	
Qy	446	TCAGGGCGAGCTCTGCTGTCCAGCGGGCCAATTCOAATCTCACACTCACGACACC	505	
Db	571	TCAGGCGCAGCTCTCTGCTGTCAAGTCGGTCAACTCAGCCCTCACCTGACAGATACG	630	
Qy	506	GAGCATGAACAACTGAGACTGATATCCGGGGCGGCTGCAGAACACGCGGGCTCCGG	565	
Db	631	GAGCACGAACAAAGTCCGACAGTAGAATGAGCAACCTGCAAGCAATCAGGCCAGTCT	690	
Qy	566	ACGCGCGCGCGCGCTCTGTGACGCGCCACACCCCCAACCCAGCACCAACGCGGCTCCATT	625	

Db 3898 AATTAAGATTTAGACATAGACGACCCAGCTCATAGATACCTTTGCAACGGATCCA 3897
Qy 3896 ATGAGTGGGCGCTTCTCTTTCTGACAGCAACAGCGCGGTCTTTAAATCAAGTCC 3955
Db 3898 GTACGGGAGATCTGTACGTTCTGACACAAACACCGCAGAAATTTATCGCCCAAGTCA 3957
Qy 3956 ACTGTGTGTGAGGACCTTGTCAAGAACTCTGAGGTGCTTGGGGACAGGTGACCAG 4015
Db 3958 CTTAGGGGGCAAGAATCTTGACTAAATAATGCAAGATCTGTGAGGGACAGGGAGCAA 4017
Qy 4016 TGCCTCCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAAGCCACACTC 4075
Db 4018 TGCCTCCGTTTGACGAGGCGAGATGTGGGATGGAGGAAGCCGCTGGAAGCCACACTC 4077
Qy 4076 ACCAATCCGAGGGTATTAAGTGGACAAAGTTTGGGCTGATCTACTTCGTGGATGGACAC 4135
Db 4078 ATGAGTCCCAAGGAATGGCAGTGTGTAAGAATGGATTAATCTACTTTGTGATGAACC 4137
Qy 4136 ATGATCAGACGCACTGATCAGATGGGATCATCTCCACCCCTGCTCGCTCTAATGATCTC 4195
Db 4138 ATGATTAGGAAGTTGACCAAAATGGAAATCATATCAACTCTTCTGGGCTTAAGATTTG 4197
Qy 4196 ACATCAGCCGCGCACTCAGCTGTGATTTCTGATGGATATTTCCAGGTAAAGCTGGAG 4255
Db 4198 ACTTCAGCCAGACCTTTAACTTGTGACACAGCATGCACATCAGCCAGGTACGCTCTGGAA 4257
Qy 4256 TGGCCACAGACTTAGCCATCAACCCAAATGGAACACTCACTTTATGTCCTCGACAAAT 4315
Db 4258 TGGCCCACTGACCTAGCCATTAACCTATGGAATCACTCAATTTATGTCCTGGATAATAT 4317
Qy 4316 GTGCTCTGCAATCTCTGAAACCCACAGGTGCGCATTTGCGCGGAGGCCATGCAAC 4375
Db 4318 GTAGTTTACAGATCACTGMAAATCGTCAAGTTGCGAATGCTGCGAGCGGCCATGCAAC 4377
Qy 4376 TGCAGGTCTCTGGCATTGACACTTCTCTGTAAGCAGGTGGCATTCACGCAACCCCTG 4435
Db 4378 TGTGAGTTCCCGAGTGG---AATATCTGTGGGGAAGCACCGGTGCGAGCAACACTG 4434
Qy 4436 GAGTCACCCACCGCTTGGCTGTTCACAAATGGGGTCTGTATATTTGCTGAGACTGAT 4495
Db 4435 GAATCAGCCACTGCCATTTGCTGTCTCAAGTGGGGTCTGTACATTAAGTAACTGAT 4494
Qy 4496 GAGAAAGAATCAACCGCATCAGCGAGTCAACACTAGTGGAGAGATCTCACTCGTTGCT 4555
Db 4495 GAGAAGAAATTAACCGGATTAAGCGAGGTCAACACAGATGGAGAAATCTCCTTAGTGGCC 4554
Qy 4556 GGGGCCCCAGTGGCTGTGACTGTAAATATGATGCCAACTGTGATTTGTTTCTGGAGAC 4615
Db 4555 GGAATACCTTCAGAGTGTGACTGCAAAATATGATGCCAACTGTGACTGTACAGAGTGA 4614
Qy 4616 GATGGTTATGCCAAGGATGCAAGTTAAATPACCCCATCTTCTTGGCTGTGTGCTGAT 4675
Db 4615 GATGGCTACGCCAAGGATGCCAACTCAGTGGCCCATCTCCTCGCTGCTTCTCCAGAT 4674
Qy 4676 GGGGAGCTCTAGCTGGCGACCTTGGGAACATCCGAATCGGTTTATCGGAAGAACAAG 4735
Db 4675 GGTACACTGTATTTGAGATCTAGGGAATATCCGATCGGGCTGTGTCAAGAATAAG 4734
Qy 4736 CCTTTCTCAACACCCAGACATGTATGAGCTGTCTTCAACATTCACAGGAGCTCTAT 4795
Db 4735 CCTTTACTTACTATGAACTTCTATGAGTTGGCTCTCCAATGATCAAGNACTTAC 4794
Qy 4796 CTGTTTGAACACCGCAAGCACTGTACACCCAAAGCCGTGCCACAGAGACTACCTG 4855
Db 4795 ATCTTTGACATCAATGGTACTCAACCAATATCTGTAAGTTTACTGCTGATGATTACCTT 4854
Qy 4856 TACAATTTCACTTACCTGGGAGCGGCATCACACTCATCACAGCAACATGGCAAC 4915
Db 4855 TACAATTTAGCTACAGCAATGCAATGATATTTACTGTGTGACACAGCAATGGCAAC 4914
Qy 4916 ATGGTAAATGTCGCGAGACTCTACTGGGATGCCCTCTGGTGGTGTCCAGATGGC 4975

Db 4915 ACCTTTGAATTTAGACGGGACCCAAATCGCATGCCAGTTTCGAGTGGTGTCTCTCTGATAC 4974
Qy 4976 CAGGTGTACTGGGTACCATGGGACCAACAGTGCACCTCAAGAGTGTGACCAACAAGGA 5035
Db 4975 CAAGTGTATGTTTACCAATAGGAACAATGGATGATCTTTGAAAGGCATGACTGCTCAAGGA 5034
Qy 5036 CACGAGTTGGCCATGATGACATACCATGGCAATCCGGCCTTCTGGCAACCAAAAGCAAT 5095
Db 5035 CTGGAATTAGTTTGTTTTACTTTACCATGGCAATAGTGGCCTTTTAGCCACTAAAGATGAT 5094
Qy 5096 GAAACCGGATGGACAAACATTTTATGATACGACAGCTTTTGGCCGCTGACAAATGTGACC 5155
Db 5095 GAAACTGGATGACAAACGTTTTTTGACTATGACAGTGAAGTCTGCTGACAAATGTTAGC 5154
Qy 5156 TTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTATACAGACAGTTCASTGATGTCCAG 5215
Db 5155 TTCCAACTGGAGTGGTCAAAACCTGCATGGGACATGGCAAGGCTTATCACAGTGGAC 5214
Qy 5216 GTAGAGACTCCAGC---AAGGATGATGTCAACATAACCAACCACTGTCTGCCCTCAGGC 5272
Db 5215 ATTGAGTATCTAGCCGAGAGAGATGTGAGCATCACTTCAATCTGTCTCTCGATCAT 5274
Qy 5273 GCCTTCTACACTGCTGCAAGACCAAGTCGGGAACAGCTACTACATCGGGCCGATGGC 5332
Db 5275 TCTTTCTACCACTGGTTCAAGATCAGTTAAAGAAACAGCTACCAAGTTGGTTATGACGGC 5334
Qy 5333 TCCTTGGCGCTGCTGCGCCAAACGGCATGGAGTGGCGCTGCAGACTGAGCCCACTG 5392
Db 5335 TCCCTCAGAAATTAATCTAGCCAGTGGCTTGACTCACACTCAAAACAGAGCCGACGTT 5394
Qy 5393 CTGCTGTCACCGCTCAACCCACCGTGGGCAAGAGGAATGTCAAGCTGCCCTCCATCGACAAC 5452
Db 5395 CTGGCTGCACCGCTAAATCCGAGGGTTGCCAAAGAAACATGACTTTTSCCTGGCGAGAAC 5454
Qy 5453 GGCCTCAACTGTGTGAGTGGCGCAGCGCAAGAGAGAGGTTCGGGGCCAGGTCACTGTC 5512
Db 5455 GGTCAAACTTTGGTGGAAATGGAGATTCCGAAAGAGCAAGCCCAAGGAAAGTCAATGTC 5514
Qy 5513 TTTGGGCGCGCTGCGGGTGCACAAACCGAAATCTCTATCTGTGAGCTTTGATCGCGTA 5572
Db 5515 TTTGGCCGCAAGCTCAGGGTTAAATGGCAGAAACCTCTTTTCACTGATTTGATCGAACA 5574
Qy 5573 ACACGCAAGAGAGATCTATGATGACCAACCGCAAGTTTCACTTTCGGAATTTCTGTACGAC 5632
Db 5575 ACAAAGACAGAAAGATCTATGACACCAACCGTAAATTTCTACTGAGGATCGCTTACGAC 5634
Qy 5633 CAGGCGGCGCGCCAGCCTCTGTGTCACCCAGCAGCAGCTGAATGTTGTCACCTGACA 5692
Db 5635 ACGTCTGGGACCCGACTCTCTGCGTGCACAGCAGCAAGCTGATGGCCGTCAATGTCACC 5694
Qy 5693 TACTCCCTTGGGGTTTACATTTGCTGGCATCCAGAGGGGCATCATGCTCTGAAAGAAATGAAA 5752
Db 5695 TATTCATCCACAGGTCAAATTTGCCAGCATCCAGCGAGGCCACCTAGCGAGAAAGTAGAT 5754
Qy 5753 TACGACAGGCGGGCCGATCATCATCCAGGATCTTCGCTGATGGGAAGACATGAGCTAC 5812
Db 5755 TATGCGACAGGGGAGGATCGTGTCTCGGCTTTTGTGATGTTGTAACATGAGGTTAC 5814
Qy 5813 ACATATTTAGAGAAGTCCATGTTGCTGCTACTACACAGCCAGAGCAGTATATCTTTGAG 5872
Db 5815 ACATATTTAGAAAGTCCATGTTCTTCTGCTTATAGCCAGCGGAGTATCATCTTCCAA 5874
Qy 5873 TTCGACAGAAATGACCGCTCTCTTCTGTGAGTATGCTCCAAAGTGGCGCGGCGGACACTA 5932
Db 5875 TACGATATGTGGGACCGCTCTCTGCCATCACCATGCCAGTGTGGGTTCGCCACACCATG 5934
Qy 5933 GAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCTCGAGGGGAATGCC 5992
Db 5935 CAGACCATCCGATCCATTTGGCTTCTACCGCAACATATACAACCCCGGAAAGCAACGCC 5994
Qy 5993 TCAGTCAATAGGACTTCACTGAGGATGGGACCTCTCTTCACTTCTACCTTCTACCTGGGCACT 6052
Db 5995 TCCATCATCGGACTACACAGAGAGGGGCTGCTTCTTACAAACAGCTTTCTTGGGTACA 6054

QY 6053 GGCGCAGGTGATATACAAAGTATGGCAAACTGTCAAAGCTGCAGAGAGCGTCTATGAC 6112
 Db 6055 AGTCGGAGGCTTATTCANAATACAGAGGCGAGCTAGGCTCTCAGAAATTTATATGAT 6114
 QY 6113 ACCACCAAGTCTAGTTTCACTATGACAGAGCGGAGGATGCTGAAGACCAATCAACCTA 6172
 Db 6115 AGCACAAGAGTCTAGTTTCACTATGATGAACACAGCAGAGTCTGAAGACAGTAACCTC 6174
 QY 6173 CAGAATCAGAGGCTTCACTGCACCATCCGCTACCGTCAGATTGGGCGCCCTGATTGACGA 6232
 Db 6175 CAGAGTATGGTTTATTGCAACATAGATACAGGCAATTTGGTCCCTGATTGACAG 6234
 QY 6233 CAGATCTTCCGCTTCACTAGGAGGATGCTCAACGCCCGTTTGAATACCAATATGAC 6292
 Db 6235 CAGATTTTCCGCTTATAGTGAAGATGGATGGTAAATGCAAGATTTGACTATAGCTATGAC 6294
 QY 6293 AACAGCTTCCGGTGACCAATGCAAGCTGCTGATCAAGAGACCCAGCTGCCATGAT 6352
 Db 6295 AACAGCTTTCGAGTGACCAATGCAAGCTGCTGATCAAGAGACCCAGCTGCCATGAT 6354
 QY 6353 CTCTATCGCTATGATGATGTGTCAGGCAAGACAGAGCAGTTTGGGAAAGTTTGGTGTCA 6412
 Db 6355 CTGTATCAGTTTATGATGATTTCTGGCAAGTTTGGAGTTTGGAGTTATA 6414
 QY 6413 TACTATGACATTAACAGATATCAACAGCTGTCTATGACCCACACCAAGCAATTTGAT 6472
 Db 6415 TATTATGATTAACCAAGATCAATTTACAGCTGTAAATGACCTATACGAGCAGCTTTGAT 6474
 QY 6473 GCATATGCGAGGATGAAGGAGTGCAGTATGATCTTCGCTGCTCATGTACTGATG 6532
 Db 6475 GCTATGCGGCTATCAAGGAGATCAATATGAGATTTCAAGTCTCATGTACTGAT 6534
 QY 6533 ACCGTCCAGTATGATAACATGGGCGAGTAGTGAAGAGAGCTGAAGGTAGGACCTTAC 6592
 Db 6535 ACAATTCAATGATATACATGCTGGGTAAACCAAGAGAGATTAATAATAGGGCCCTTT 6594
 QY 6593 GCGAATACCACTCGCTACTCTATGATGATGCTGAGCGGCGAGCTGCAGACAGTCTCC 6652
 Db 6595 GCGAATACCACTCGCTACTCTATGATGATGCTGAGCGGCGAGCTGCAGACAGTCTCC 6654
 QY 6653 ATCAATGACAGGCACTCTGCGCTACAGCTAGCAGCTCAATGGGAACTCTGCACTTACTG 6712
 Db 6655 CTCAATGAAAGATAATGTGCGGTACAATACGATCTGAATGGAACCTCCATTTACTG 6714
 QY 6713 AGCCTGGGACAGTGCAGCGCTCACCACTACGCTATGACATCGGACCGCATCACT 6772
 Db 6715 AACCAAGTAAACAGTGGCGCTCTGACACCCCTTTCGCTATGACCTGGAGACAGAACTCA 6774
 QY 6773 CGGTGGGTGACGTGCAATACAGATGGATGAGGATGGCTTCTGAGGAGCGGGCGGT 6832
 Db 6775 CGACTGGGTGATGTTCAATATCGTTGGATGAGATGGTTCTACGTCAAGGGGACAG 6834
 QY 6833 GATATCTTGGATACAACTCAGCTGGCTGCTCATCAAGCGCTTACAAACCGGCTGCGAGC 6892
 Db 6835 GAAATCTTTGAATATAGCTCAAGGGGCTTCAACTCGAGTTTACAGTAAAGGCACTGGC 6894
 QY 6893 TGGAGTGTCAAGTACCGCTACGATGGCTGCGGCGCGCTGCTCAGAGAGAGCAGCCAC 6952
 Db 6895 TGGACAGTATCTACCGTTTATGACCGCTGGGAGGCGGTGTTTCTAGCAAAACCACTCTA 6954
 QY 6953 AGCCACCACTTGCAGTTTCTTATGACAGCTCAACCAACCCCAAGGTCAACCCACTG 7012
 Db 6955 GGACAGCACTGCACTTTTATGCTGCTTAACTATATCCCACTAGGATTTACTCATGTC 7014
 QY 7013 TACAACCACTCCAGCTCTGAGATCACTCCCTCTCTACGACTTGAAGAGACACTCTTT 7072
 Db 7015 TACAACCACTTCCAGTTTACAGAAATTTACCTCCCTGTTATGATCTCAAGGACATCTTT 7074
 QY 7073 GCCATGAGCTGAGCAGTGGTGTAGTTTATCATAGCTTTGTGACAACTCGGACCCCT 7132
 Db 7075 GCCATGGAATCAGCAGTGGGATGAATTTCTATATTGCACTCGGATAACACAGGGAACCA 7134

QY 7133 CTTCTGTCTTTAGTGGAAACAGGTTTGATGATCAAGCAAACTCTGTACAGAGCTATGGG 7192
 Db 7135 CTGCTGTGTTCACTAGTAATGGCTTATGCTGAACAGATTCAGTACACTGCATATGGG 7194
 QY 7193 GAGATCTACATGGATACCAACCCCACTTTTCAGATCATATAGGCTACCACTGGTGGCTC 7252
 Db 7195 GAAATCTATTTTGACTCTAATATGACTTTCACTGTAATTTGATTTCACTGGTGGCTC 7254
 QY 7253 TATGATCCACTCACCAAGCTTGTCCATCGGCGGCGAGATTTATGATGCTGTGGCGGA 7312
 Db 7255 TATGACCCACTCACCAAAATTAATCCACTTTGGAGAAAGATTTATGACATTTTGGCAG 7314
 QY 7313 CGTGGATAGCCAGACAGGCTGTGGAAGCACCTTAGTAGCAGCAAGCTCATGCTC 7372
 Db 7315 CGTGGACCAACACTGACATAGAAATCTGGA---AAGAATTGGGAAGACCCAGCTCCT 7371
 QY 7373 TTTAATCTCTATATGTTCAAAAACAACACCCCTCAGCAAACTCCAGGACATCAAGTGC 7432
 Db 7375 TTTAATCTTGTACATGTTTAGGAATAACAACCTGCAAGCAAAATCCATGACGTGAAGAT 7431
 QY 7433 TTCATGACAGATGTTAAGCTGCTGCTCACTTTGGATTCAGCTTACAGCTACACAGCTGATC 7492
 Db 7435 TACATCAGATGTTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7491
 QY 7493 CCTGTTATCCAAACACAGACATGATGCTGCAAGCACTCTACAGAGCTCATCCACACA 7552
 Db 7495 CCTGATTCCTGTTCCAAATTTGATTTAAGCAAGCTTCTTACGA-----A 7539
 QY 7533 CAGATGAAACCGAGAGTGGGACAAACAGCAAGTCTATCTCGGGGTACAGTGTGAAGTA 7612
 Db 7535 CTTGTGAAGAGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7599
 QY 7613 CAGAGCAGCTCAAGGCTTTGTCACCTTAGAACGTTTGCACAGCTCTATGCTCCACA 7672
 Db 7600 GCGCGGAGGCAAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7659
 QY 7673 ATCACCAGCTGCCAGAGGCTTCAAGAACCAAGAAAGTTTGCATCCAGCGCTCAGTCTTT 7732
 Db 7660 GCGCGGCGGCT---GGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7716
 QY 7733 GCGAAGGGGTCAAGTTTGCCTTGAAGATGGCGGAGTGCAGACAGACATCATCAGTGTG 7792
 Db 7717 GCGAAGGGGTCAAGTTTGCCTTGAAGATGGCGGAGTGCAGACAGACATCATCAGTGTG 7776
 QY 7793 GCGAATGAGGATGGCGAAGGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7852
 Db 7777 GCGAAGGAGTGCATCAAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7836
 QY 7853 CTGACCTTCACTATGATGGGTGGATGCCATTTACTTTGTGAACACAGGACCTTCAAG 7912
 Db 7837 CTGACCTTCACTATGATGGGTGGATGCCATTTACTTTGTGAACACAGGACCTTCAAG 7896
 QY 7913 GGTGACCTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 7972
 Db 7897 AGGACCTTGGGACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7956
 QY 7973 GTCATGTTGCCAGATCAACACAGTACTTAAATGGCAGGACTAGACGCTACACAGATC 8032
 Db 7957 GTGACGTTGCGAGTCCACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8016
 QY 8033 CAGCTCAGTACGGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8092
 Db 8017 GAGATGAGTTCGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8076
 QY 8093 AAGGACGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8152
 Db 8077 AAGGCGGCTATCTTGGAGCAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8136
 QY 8153 CAGCAGAGACTTGGGAAAGGGAGGAGGCTTGGGCGCTTGGACAGAGGGGGAGAGCAG 8212
 Db 8137 CAGCAGCGCTGCGGAGCGGAGGAGGCGGCTTGGAGCGGAGGAGGAGGAGGAGGAGGAG 8196
 QY 8213 CAGGTGCTGAGCACAGGCGGCTGCAAGGCTTACAGCGCTTTTTCGTGATCTCTGTCTG 8272

Db 8197 CAGCTGTAGCGCGCAAGGTGAGGCTACGCGGTACTACGTACTCTCGGTGAG 8256
Qy 8273 CAGTACCCAGACTGTGAGCAGCGCCAAACACATCCACTTCATGACAGAGCGAGATG 8332
Db 8257 CAGTACCCAGCTGTGGCGAGCGCCAAACACATCCACTTCATGAGCGAGATC 8316
Qy 8333 GCGCGAGGTGAC 8345
Db 8317 GGCAGGAGGTAAAC 8329

RESULT 7

US-10-038-854-39

; Sequence 39, Application US/10038854

; Publication No. US20040022781A1

; GENERAL INFORMATION:

; APPLICANT: Spytex, Kimberly A

; APPLICANT: Li, Li

; APPLICANT: Wolenc, Adam R

; APPLICANT: Vernet, Corine

; APPLICANT: Eisen, Andrew J

; APPLICANT: Liu, Xiaohong

; APPLICANT: Malvankar, Uriel M

; APPLICANT: Shimkets, Richard A

; APPLICANT: Tchernev, Velizar

; APPLICANT: Spaderna, Steven K

; APPLICANT: Gorman, Linda

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Patturajan, Meera

; APPLICANT: Gusev, Vladimir Y

; APPLICANT: Gangolli, Esha A

; APPLICANT: Guo, Xiaojia S

; APPLICANT: Shenoy, Suresh G

; APPLICANT: Rastelli, Luca

; APPLICANT: Casman, Stacie J

; APPLICANT: Boldog, Ferenc

; APPLICANT: Burgess, Catherine E

; APPLICANT: Edinger, Shlomit R

; APPLICANT: Ellerman, Karen

; APPLICANT: Gunther, Erik

; APPLICANT: Smithson, Glenda

; APPLICANT: Millet, Isabelle

; APPLICANT: MacDougall, John R

; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

; FILE REFERENCE: 21402-230

; CURRENT APPLICATION NUMBER: US/10/038,854

; CURRENT FILING DATE: 2003-01-22

; PRIOR APPLICATION NUMBER: 60/258,928

; PRIOR FILING DATE: 2000-12-29

; PRIOR APPLICATION NUMBER: 60/259,415

; PRIOR FILING DATE: 2001-01-02

; PRIOR APPLICATION NUMBER: 60/259,785

; PRIOR FILING DATE: 2001-01-04

; PRIOR APPLICATION NUMBER: 60/269,814

; PRIOR FILING DATE: 2001-02-20

; PRIOR APPLICATION NUMBER: 60/279,832

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,833

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/279,863

; PRIOR FILING DATE: 2001-03-29

; PRIOR APPLICATION NUMBER: 60/283,889

; PRIOR FILING DATE: 2001-04-13

; PRIOR APPLICATION NUMBER: 60/284,447

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 60/286,683

; PRIOR FILING DATE: 2001-04-25

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 411

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 39

; LENGTH: 8473

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-038-854-39
Query Match 41.0%; Score 3422; DB 17; Length 8473;
Best Local Similarity 67.1%; Pred. No. 0;
Matches 5027; Conservative 0; Mismatches 2370; Indels 93; Gaps 9;
Qy 880 CGGGCACTTCCTCTTCAAGCCTGG---AGGCACTTCCCGCTCTTCTGACACCATCACC 936
Db 725 CAGGCATTTCTTATTAACAAACAGGAACAGGTACAAAGCCACTGTTTCACTACTGCAACCCC 784
Qy 937 AGGGTACCCACTGACGTCCAGCACAGTGTACTCTCTCCGCCCGCCAGCCCTCCCGCGAG 996
Db 785 AGGATACACATGGCATCTGGCTCTGTTTATTTACCACTACTCTCGGCCACTACTCTAGAAA 844
Qy 997 GACCTTCGCCCGCGCGCTTTAACTCAAGAGCCCTCAAGTACTGTACTGTGAAGTG 1056
Db 845 CACCCTATCAAGAAGTGTCTTTAAATCAAGAAGTCTTCAAAGTACTGTAGTGGAAATG 904
Qy 1057 CGCAGCCCTGAGCGCATCGTCTATCTCAGCCACTCTGGTCTATCTCTCTCTGCTGGATCTTTGT 1116
Db 905 CACTGCATCTGTGCGGTAGGGTCTCGTGCTCTCTGCAATACTCTCTGTTCTTATTTAT 964
Qy 1117 GGCATGCACTGTTTGGCTTAACTGGACCTCTGAGCCGATGGAGGGCGAGATGTATGA 1176
Db 965 AGCAATGCATCTCTTTGGCTCACTGGCAGCTACAGCAGACTGAAAAATGACACATTTGA 1024
Qy 1177 GATCAGGAGGACACAGCAGCAGTGGCTGTGCCAACCGACGCTCTCCCTATACCCCTC 1236
Db 1025 GA-----ATGGAAGAGTGAATTCGTATACATGCAACAAACACTGTGTCTATACCTTC 1078
Qy 1237 AGGGGCGACTGGCTTTAGAGACCCCTTGACAGAAAGGCAAGAACCCACAGAGGAAGGCC 1296
Db 1079 TGGAGACA-----ATGGAAGATTT 1096
Qy 1297 CAGTAGTTTCTTCCAGAGGACAGTTTCATAGATTCTGGAGAAATGATGGGAAGCG 1356
Db 1097 AGGTGGATTTAGCAAGAAATAACACATAGATTTCGGAGAACTTGATATTGGCCGAG 1156
Qy 1357 AGCCTCCCAAGATTCTCTCTGGCATCTTCTGGAGATCTCAAGTGTTCATAGACATCC 1416
Db 1157 AGCAATTCAAGAGATTCTCTCCGGATCTTCTGGAGATCACAGCTCTTCAITGATCAGCC 1216
Qy 1417 TGTGATCTGAAATTCATGTGTCTCTGGGAAAGGACGCTCTGTTGGCATTTATGGCAG 1476
Db 1217 ACAGTTTCTTAAATTCATATCTCTTTCAGAAAGGATGCAATTGATGGATATATGGCG 1276
Qy 1477 AAAAGGCTCTCTCTTCCACATACACAGTTTGTGACTTTCTGGAGCTGCTGGATGGCAGG 1536
Db 1277 GAAAGCTTACCGCTTCCATCTACTAGTATGACTTCTGGAGCTCTCTGGATGGCAGG 1336
Qy 1537 GCTCTTAAACCCAGGAGCGCGAGCTTAGAGGGGACCCCGCCAGTCTCGGGGAACCTGT 1596
Db 1337 GCTGATTCGACAGAGCAGCGAGCGCTGCTTTCAGACGAGAGAGCGCGCGCGCAGCGAG 1396
Qy 1597 GCGCCCTCCAGCCATGAGCAGGCTTTCATCAGTATTTGGATTCAGAAATCTGGCATT 1656
Db 1397 ATCCCTCAGCTTTCATGAGCGCGCTTATCCAGTACTTGGATTCTGGAATCTGGCATCT 1456
Qy 1657 GCGCTTTTACATGACGGAAGAGTTCAGAAAGTGTCTTCTCTTCTCACCCTGCCATTGA 1716
Db 1457 GCGTTTTTATGATGGGAAATGACAGCAGCGTGTCTTTAATACCATTTGTTATAGA 1516
Qy 1717 GTCGTGATTAACCTCCCGCAGCACTGTATGGGAATGGTGAATGCTCTCTGGACCTG 1776
Db 1517 GTCTGTGGTGAATGTCCCGGAAATGCAATGGAAATGGAGAAATGCGGTTCTTGGAACTG 1576
Qy 1777 CCACCTGCTTCTGGGTTTCTCGGGCCCGAGCTGTGGCAGAGCCCTCTCCCGCTGCTG 1836
Db 1577 CCATTTGTTTCCAGGATTTCTGGGTCCGATTTTCAGAGCCCGCTCTCCAGTGTGTA 1636
Qy 1837 TAGCGGAAATGGCCAATACATGAAAGGCGAGATGCTTGTGCCACAGTGGCTGGAAGGCGC 1896

Db 1637 TAGTGGCAACGGGAGTACTCCAGAGCGCGCTGCTGTTTTCAGCGGCTGGAGGAC 1696
 Qy 1897 TGAGTGCATGTGCCCCAACCAAGTGTATCGATGTGGCTGCGAGCAACATGGACCTG 1956
 Db 1697 CGAGTGTGATGTGCGGATACCCAGTGTATGACCCACAGTGTGGGGTGTGGGATTTG 1756
 Qy 1957 CATCACGGGACCTGATCTGCAACCTGCTCAAGGGCGAGCTGTGAGGAGTGA 2016
 Db 1757 TATCATGGGCTCTTGTGCTTTCGAACCTCAGGATCAAGGAAAGTTGTGAAGACTGA 1816
 Qy 2017 CTGCATGGACCCACATGTTTCAGCGCGGGTGTCTGCTGAGAGGGAATGCCATTGCTT 2076
 Db 1817 CTGTATAGACCTGGGTGTTCTAATCATGTGTGTGTATCCAGGGGAATGCACTGCAG 1876
 Qy 2077 TGTGGGATGGGAGGACCAACTGCGAGACCCCGAGGGCCACATGTGTAGACAGTGTTC 2136
 Db 1877 TCCAGGATGGGAGGTAGCAATTTGAAATACTGAAGACCATGTGTCCAGACCAAGTCTC 1936
 Qy 2137 AGGCCAGGAACCTTCTCTCCGGACACCGGGCTTTGCGAGCTGTGACCCAAAGCTGGACTGG 2196
 Db 1937 CGGCCAGGAACGTATCTTCAAGAAAGTGTCTCTGCACTGTGACCTTAACTGGACTGG 1996
 Qy 2197 ACAGACTGTTCATCGAGATCTGTGCTGCCGACTGTGTGGCCATGGCGTGTGCTAGG 2256
 Db 1997 CCCAGACTGTCAAAAGAAATATGTTCTGTGGACTGTGGCTCACAGGGGTTTGTGATGG 2056
 Qy 2257 GGGACCTGCGCTGCGAGTGGCTGTGGATGGGGGAGCTGCGACCGGGGCTGGCA 2316
 Db 2057 GGGACCTGCTGCTGTGAAAGAGCTGTGACGGGCGGAGCTGTAACTAGAGAGCTGGCA 2116
 Qy 2317 CCCCGCTGTGCCAGCATGGGACCTGCGCGAGCGCAAGTGGAGTGGAGCCCTGGCTG 2376
 Db 2117 CCCCGCTGTGCCAGCACGGGACCTGCAAGGATGGCAAGTGTGAATGAGCCAGGGCTG 2176
 Qy 2377 GAATGGCGAACAATGCAACCTCGTCTACTATCTGGATAGGTTAGT----- 2421
 Db 2177 GAATGGAGAGACTGCACTATCGCTACTATTTGGATTAAGTATGTTAAAGACAAGATAGG 2236
 Qy 2422 ---TAAAGAGGTTGCTCGGTTGTCAATGGCAACCGGAGATGTACCTTATAGACTGAA 2478
 Db 2237 ATATAAGAGGTTGCTCTGCTGTGCAACAGCATGGAAGATGTACCTTGGACCAAA 2296
 Qy 2479 TGGTTGCACTGCTGTGCCAGTGGCTGGAGGAGCTGGCTGTGACACTTCCATGGA 2538
 Db 2297 TGGCGGACATTTGTGTGCGAGCTGTGATGGAGGAGGAGGCTGTGACGTAGCCATGGA 2356
 Qy 2539 GACTGCTGCGGTGACAGCAAGCAATGATGGAGTGGCTGTGGACTGCAATGGACCC 2598
 Db 2357 GACTCTTTTGCAGATAGAGAGCAATGAAGGATGAGTCAATGATGCAATGATGATCC 2416
 Qy 2599 TGACTGTGCTCCAGCGCCCTGTGCCATATCAACCGGCTGTGCTTGGCTCCCTAACCC 2658
 Db 2417 CGATTGTGCTCAGAGTTCTCTGCCAGATCAGCCCTATTGTGCGGGAGTGGCGGATCC 2476
 Qy 2659 TCTGGACATCATCAGAGACAGAGTCCCTGTGTACAGACAGAACCTACACTCTCTTA 2718
 Db 2477 TCAGGACATCAATAGCCAAAGCCCTTCAATCGCTTCTCAGCAAGTGGCCAAATCCCTTTA 2536
 Qy 2719 TGACCGCATCAAGTTCTCTGTGGGAGGAGCAGCAGCAATATCCCGGGGAGAACCC 2778
 Db 2537 TGATCGAATCAGTTTCTTATAGATCTGATAGACCCATGTTATACCTGGAGAAAGTCC 2596
 Qy 2779 CTTTGTAGGGGAGCATGCTGTGTATTCGTGGCCAGTGTATGACATCAGATGGAACCC 2838
 Db 2597 TTTCAATAAGAGCCCTGTGATCTGTGATCAGAGGCAAGTACTGACTGCTGATGGAACCTC 2656
 Qy 2839 CTTGGTTGGTGAACATCAGTTTGTGTAATCAATCAATCCCTCTTTGGATATACATCAGCAG 2898
 Db 2657 ACTTATGGAGTAATGTCCTGTTTTTCCATTACCCAGATATGATATACTATTACCCG 2716
 Qy 2899 GCAAGATGGGAGCTTGTGATGTAATAATGGCGGATCTCCATCATCTCTCGGTTGCGAG 2958

Db 2717 CCAGGACGAATGTTGACTTGTGTGGCAAAATGTTGGGCGCTCTCTAACTTTGTTATTTGA 2776
 Qy 2959 CGGGGACCTTTTCATACACAGGAGCACCCCTGTGTGCTGCCATGGGATCGCTTTTGT 3018
 Db 2777 ACGATCCCATTTCTCTACTCAGTATCATCTGTGTGATTCATGGAATGCTCTTTATGT 2836
 Qy 3019 CATGGAACCATCATCATGAGACATGAGGAGAAATGAAITCCCAGCTGTGACCTGAGCAA 3078
 Db 2837 GATGATACCTTAGTCAATGGAGAAAGAGAAATGACATTTCCAGCTGTGATCTGAGTGG 2896
 Qy 3079 TTTTGGCGGCGCCAAACCGAGTGTCTCTCATCCCCACTGAGCTCTTCGCCAGCTCTG 3138
 Db 2897 ATTCTGTAGGCGCAATCCCATATTGTGTGATCATCACTTTATCCACCTTTTTCAGATCTC 2956
 Qy 3139 TGCAGAGAAAGGCGCCCATTTGTCGGGAAATTCAGGCTTTTCAGGAGGAAATCTCTATCTC 3198
 Db 2957 TCCTGAAGACAGTCCCATCATTTCCGGAACACAGTACTCCACGAGAAATCAATTC 3016
 Qy 3199 TGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGAGCCCTGGCTTACAAATCTGTCT 3258
 Db 3017 AGGAACAGATTTGAAACTCTCTACTTCACTTCACTTCCAGAGCTGCAAGGTATTAAGTCAGTTCT 3076
 Qy 3259 GAGGATCAGCTTCAACCCAGCCAGCATCCCTTCAACCTCATGAAGGTGCACTCATGCT 3318
 Db 3077 CAAGATCACCATGACCCAGTCTATTATTCATTTAATTAATGAAGTTTCTCTTATGTT 3136
 Qy 3319 AGCGTGGAGGCGCCCTCTTCAGGAAGTGTCTGCTGCAAGCCCGAGCCAGCTGTCTATTA 3378
 Db 3137 AGCTGTAGTGAAGAGCTTCTTCCAAAAGTGTGTTCTTCCCTCACCACAACTTGGCTATAC 3196
 Qy 3379 TTTCAATTTGGAGAACAGACAGCTCTACACAGAGAGTGTGTTGGGCTTTTCAGAGCTT 3438
 Db 3197 TTTCAATTTGGGATAAAAACAGATGATATAATCAGAAAAGTCTATGTTCTATCTGAAGTGT 3256
 Qy 3439 TGTTCCTGGCTTATGAATATGAATCTGCGCCAGATCTAATCTGTGGGAAAAAAGAAC 3498
 Db 3257 TGTGTAGTTGGATATGAGTATGAGTCTGTTTGGACCTGACTGTGTGGGAAAGAGGAC 3316
 Qy 3499 AACAGTGTGCGGGCTATGAAATTTGAGCGTCCAAGCTTGGAGGATGGAGCCCTAGACAA 3558
 Db 3317 TGCCATTTGCGAGGCTATGAATTTGGATGCGTCCACATGCGTGGCTGGAATTAGATAA 3376
 Qy 3559 ACATCATCCCTCAACATTCAGAGTGTATCTCTCACAAGAGGAATGGGAGAACCAAGTT 3618
 Db 3377 ACATCAGTGTGATGTACAGAACGTTATCTGTACAAGGAAACCGGGAACCAAGTT 3436
 Qy 3619 TGTGTCTCAGAGGCTCTGTCTATTGGAGCATCATGGCAATGGGCGCGGAGAACAT 3678
 Db 3437 CATCTCCAGCAGCCTCCAGTCTGAGTAGCATCATGGCAATGGGCGAAGCGCAGCAT 3496
 Qy 3679 CTCCTGCCAGCTGCAACGCTTGTGAGGCAACAGCTCTGCGCCCGGAGTGGCCCT 3738
 Db 3497 TTCTTGGCGGAGTGTCAATGCTCAAGTGTGATGGAAGTTACTGCGCCCGAGTGGCT 3556
 Qy 3739 CACTGTGCTCTGACGGGAGCTCTATGTTGGTGAITTTCAACTACATTTAGAGGATCTT 3798
 Db 3557 AGCTTGTGGATCGATGGCACTCTGTAGTAGGCAATTTCACTACGTGGCGGATATT 3616
 Qy 3799 CCCCTCTGGAAATGTCCAAACATCTTAGAGCTAGGAAATGAAGTTTCAGACATAGTCA 3858
 Db 3617 CCGTTCTGGAATGTAAACAGTGTCTAGAACTAAGAAATTAAGATTTTAGACATAGCAG 3676
 Qy 3859 CAGTCCACACACAAATACTACCTGGCCACAGACCCCATGAGTGGGCGCTCTTCTCTTC 3918
 Db 3677 CAACCCAGCTCATAGATACCTTGTCAACCGATCCAGTACGGAGATCTGTACGTTTC 3736
 Qy 3919 TGACAGCAACCGCGGGTCTTTTAAATCAATCCATCTGTGTGTGAGGAGCTTGT 3978
 Db 3737 TGACACAAACACCGCGAGATTTTATCGCCCAAGTCACTTACGGGGGCAAAAGACTTGC 3796
 Qy 3979 CAAGAACTCTCAGTGTGTTGGGGACAGGTGACAGTGCCTCCCTTTTGTATGACACTCG 4038
 Db 3797 TAAAAATCGAAGTCTCTCGGACAGGACAGGGAGCAATGCTTCCGTTTTCAGAGGCGAG 3856

6014	Db		CAATTAGATACAGGCAAAATGGTCCCTTGATTGACAGGCAGATTTCGCGCTTTAGTGAAGA	6073
6256	Qy		AGGCATGGTCAACGCCCGTTTGGACTACAACATATGACAACAGCTTCGCGGTGACCAAGCAT	6315
6074	Db		TGGGATGGTAAATGCAAGATTTGACTATAGCTATGACAAACAGCTTCCAGTGCACAGCAT	6133
6316	Qy		GCAGGCTGTGATCAACAGAGACCCACATGCCCATTTGATCTCTATCGTATGATGATGTTC	6375
6134	Db		GCAGGGTGTGATCAATGAAACGCCACATGCTATTTGATCTGTATCAGTTTGTATGACATTTT	6193
6376	Qy		AGGCAAGACAGAGCAGTTTCGGAGTTTGGTGTCATTTTATCTATGACATTAACACAGATCAT	6435
6194	Db		TGGCAAAAGTTGAGCAGTTTGGAAAGTTTGAGTTTATATTTATGATTTAAACAGATCAT	6253
6436	Qy		CACCACAGCTGTCTATGACCCACACCAAGCATTTTGTATGTCATATGGCAGGATGAAGAAAGT	6495
6254	Db		TTCTACAGCTGTAATGACCTATACGAAGCACTTTGATGCTCATGGCCGTATCAAGAGAT	6313
6496	Qy		GCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATGACCGTCCAGTATGATTAACATGGG	6555
6314	Db		TCAATATGAGATATTTACGTCGCTCATGTACTTGGATTACAATTCAGTATGATTAACATGGG	6373
6556	Qy		GCAGTAGTGAAGAAGAGAGCTGAAGGTAGACCCCTAGCCCAATACCACTCGCTACTCTCTA	6615
6374	Db		TCGGTTAACNAGAGAGATTAATAATAGGGCCCTTTGCCAAACCAACCAATATGCTTA	6433
6616	Qy		TGAGTATGATGCTGACCGCGAGCTGCAGACAGTCTCCATCAATGACAAGCCATCTTGGCG	6675
6434	Db		TGAATATGATGTTGATGGACAGCTTCCAAAACAGTTTACCTCAATGAAAGATAATTTGGCG	6493
6676	Qy		CTACAGCTACGACTCAATGGGACCTGCACTTACTGAGCCCTGGGAACAGTGCACGGCT	6735
6494	Db		GTAACAATCGATCGAATGGAAACCTCCATTTACTGAAACCAAGTAACAGTTCGGGTCT	6553
6736	Qy		CACACCACTACGSTATGACATCCGCGACCGCATCACTCGGCTGGTGACGTGCAATACAA	6795
6554	Db		GACACCCCTTCGCTATGACCTGGAGACAGAACTCACTCGACTGGTGATGTTCAATATCG	6613
6796	Qy		GATGGATGAGGATGGCTTCTTGAGGACAGCGGGCGGTGATATCTTTGAGTACAACCTCAGC	6855
6614	Db		GTTTGGATGAAGATGGTTTCCCTACGTCGCAAAAGGGCAGCGAAATCTTGAATATAGCTCCAA	6673
6856	Qy		TGGCCTGCTCATCAAGCGCTACAAACCGGCTGGCAGCTGGAGTGTCAAGTACCGCTACGA	6915
6674	Db		GGGCTCTTAACCTCAGTTTACAGTAAGGCAGTGGCTGGACAGTGATCTACCGTTATGA	6733
6916	Qy		TGGCCTGGGGGGCGGTGTCAGACAGAGAGAGCCACAGGCCACCACTGCAGTCTCTCTA	6975
6734	Db		CGGCTCGGGAAGGCGTGTCTTAGCAAAACCAAGTCTAGGACAGCACCTGCAGTGTTTT	6793
6976	Qy		TGCAGACTGACCAACCCCAAGGTCAACCACCTGTACAACCACTCCAGCTCTGAGAT	7035
6794	Db		TGCTGACTTTACTTTATCCCACTAGGATTACTCATGCTCAACCAATTCGAGTTCAGNAAT	6853
7036	Qy		CACCTCCCTCTACTACGACTTTGCAAGACACCTCTTTGCCATGGAGCTGAGCAGTGGTGA	7095
6854	Db		TACCTCCCTGTATTATGATCTCCAAGGACATCTTTTGGCATGGAAATCAGCAGTGGGA	6913
7096	Qy		TGAGTTTTACATAGCTTGTACACATATCGGACCCCTCTTGTCTCTTTTGTAGTGGACAGG	7155
6914	Db		TGAATCTTATTTGCATCGGATAACAGGGACACCACTGGCTGTGTTTCAGTAGCAATGG	6973
7156	Qy		TTTTCATGATCAAGCAAAATCTGTACACAGCCTATGGGGAGATCTACATGGATACCAACC	7215
6974	Db		GCATTATGCTGAAAACAGATTCAGTACACTGTCATATGGGAAATCTATTTTGCATCTAATAT	7033
7216	Qy		CAACTTTTCAGATCATATAGGCTACCATGGTGGCTCTATGATCCACTCACCAGCTTGT	7275
7034	Db		TGACTTTTCAACTGGTAATTTGATTTTCAATGGTGGCTGTATGACCCACTCACCAAATTAAT	7093
7276	Qy		CCCATATGGGCGCGAGATTATGATGTGCTGGCCGAGCGCTGGGACTAGCCAGCACCA	7335

Db	7094	CCACTTTGGGAAAGAGATTATGACATTTTGGCAGGACCGGTGAGACAACACCTGACATAGA	7155
Qy	7336	GCTGTGAAGACACCTTAGTAGCAGCAACGTCATGCCCTTTTAAATCTCTATATGTTCAAAA	7395
Db	7154	AATCTGGA--AAAGAAATTGGGAAGACCCAGCTCCTTTTAACTGTGTACATGTTTAGGAA	7210
Qy	7396	CAACAACCCCATCAGGAACTCCGAGACATCAAGTGTCTTCATCACAGATGTTTAACAGCTG	7455
Db	7211	TACAAACCTGTGAAGCAAAATCCATGACGTGAAGATTACATCACAGATGTTTAACAGCTG	7270
Qy	7456	GCTGCTCACCTTTGGATTCCAGCTACACAACGTATCCCTGGTTATCCCAAAACAGACAT	7515
Db	7271	GCTGGTGACATTTGGTTTCCATCTGCACAATGCTATTCTCTGGATTCCCTGTTCCCAAAAT	7330
Qy	7516	GGATGCCATGGAACTCCTTACGAGCTCATCCACAGACATGAACACGAGAGTGGGA	7575
Db	7331	TGATTTTAACAGAACCTTCTTACGA-----ACTGTGAAGAGTCAGCAGTGGGA	7378
Qy	7576	CAACAGCAAGTCTATCCTCGGGTACAGTGTGAAGTACAGAAGCAGCTCAAGGCCCTTTGT	7635
Db	7379	TGATATACGCCCATCTTCGGAGTCCAGACGAAGTGGCGCGCAGGCCCAAGGCCCTTCCT	7438
Qy	7636	CACCTTAGACGGTTGACACAGCTCTATGCTCCACATCACAGCTGCCAGCAGACTCC	7695
Db	7439	GTCCTGGGAAGATGCCAGGTGTAGGTGAGCGCGCGCGGGCCGGC---GGCGCGCA	7495
Qy	7696	AAAGACCAAGAAGTTTGTCATCCAGCGCTCAGTCTTTTGGCAAGGGGTCAAGTTTGCTT	7755
Db	7496	GTCCTGCTGTGGTTGGCCACGGTCAAGTCGCTGATCGGCAAGGGGCTCATGTGCGCT	7555
Qy	7756	GAAGTAGSCGAGTGACACAGACATCATCAGTGTGGCCAAATGAGATGGCGGAAGGT	7815
Db	7556	CAGCCAGGCGCGGTGCAGACCAACGTGCTCAACATGCCACAGAGACTGCATCAAGT	7615
Qy	7816	TGCTGCCATCTTGAACCATCCCACTACCTAGAGAACCTGCATTCACCAATTGATGGGT	7875
Db	7616	GGCGGCGTGTCTCAACAACGCTTCTACCTGGAGAACCTGCATTCACCATCAGAGGCAA	7675
Qy	7876	GGATACCCATTACTTTGTGAACAGGACCTTCAGAGGTGACTCGGCATCTGGGCGCT	7935
Db	7676	GGACACGCATCTCTCATCAAGACCAACACGCCCGAGAGCGACTGGGACGCTGCGGT	7735
Qy	7936	CAGTGGGGGGCGGCGAACCCCTGGAGAATGGGGTCAAGCTCACTGTGTCCAGATCAACAC	7995
Db	7736	GACCAGCGCGCAAGCGCTTGAGAACCGGCATCAACGTGACGTGTGCGAGTCCACAC	7795
Qy	7996	AGTACTTAATGGCAGGACTAGCGCTACAGACATCCAGCTCCAGTACGGGCACCTGTG	8055
Db	7796	GGTGGTGAACGGCAGAGACCGCGAGTTTCGCGGACGTGGAGATGCAGTTCGGCGCGTGGC	7855
Qy	8056	CTTGAACACACGCTACGGGCAACGTTTGGATGAGGAAGGCAACGGTCTGTGAGCTGGC	8115
Db	7856	GCTGCAGTGTGCTACGGCATGCCATGACCTGGACAGGAGAAGCGCGCATCTCTGAGCAGGC	7915
Qy	8116	CCGSCAGAGACCGGTGCGCCAAAGCTGGGCGCGGACGACGACGACTCGCGGAGGGA	8175
Db	7916	CGCGACGCGCGGTTCGCCCGGCGCTGGGCGCGGAGCAGCGCGCTGCGACGCGCA	7975
Qy	8176	GGAAAGGCTCGGGCCTTGGACAGAGGGGGAGAACGACAGTCTGTAGCACAGGGGGGT	8235
Db	7976	GGAGGGCGCGCCTCTGGACGGAGGCGGAGAGCGGCAGCTGCTGAGCGCGCAAGGT	8035
Qy	8236	GCAGGGCTACGACGGCTTTTGTGATCTCTGTGCGACGACTACCCAGAACTGTACAGAC	8295
Db	8036	GCAGGGCTACGACGGGTACTACGTACTCTCGTGTGAGCACTACCCGAGTGTGCCGACAG	8095
Qy	8296	CGCCAAACAACCTTCATATGACAGACGAGATGGCGCGGAGGTGAC	8345
Db	8096	CGCCAAACAATCCAGTTCTCTGGGACAGACGAGATCGGACAGAGGTAAAC	8145

	: Sequence 41, Application US/10038954	
	: Publication No. US20040022781A1	
	: GENERAL INFORMATION:	
	: APPLICANT: Spytel, Kimberly A	
	: APPLICANT: Li, Li	
	: APPLICANT: Wolenc, Adam R	
	: APPLICANT: Vernet, Corine	
	: APPLICANT: Eissen, Andrew J	
	: APPLICANT: Liu, Xiaohong	
	: APPLICANT: Malyankar, Uriel M	
	: APPLICANT: Shimkets, Richard A	
	: APPLICANT: Tchernev, Velizar	
	: APPLICANT: Spaderna, Steven K	
	: APPLICANT: Gorman, Linda	
	: APPLICANT: Kekuda, Ramesh	
	: APPLICANT: Patturajan, Meera	
	: APPLICANT: Gusev, Vladimir Y	
	: APPLICANT: Gangolli, Bsha A	
	: APPLICANT: Guo, Xiaoqia S	
	: APPLICANT: Shenoy, Suresh G	
	: APPLICANT: Rastelli, Luca	
	: APPLICANT: Casman, Stacie J	
	: APPLICANT: Boldog, Ferenc	
	: APPLICANT: Burgess, Catherine E	
	: APPLICANT: Edinger, Shalomit R	
	: APPLICANT: Ellerman, Karen	
	: APPLICANT: Gunther, Erik	
	: APPLICANT: Smithson, Glenda	
	: APPLICANT: Millet, Isabelle	
	: APPLICANT: MacDougall, John R	
	: TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same	
	: FILE REFERENCE: 21402-230	
	: CURRENT APPLICATION NUMBER: US/10/038,854	
	: CURRENT FILING DATE: 2003-01-22	
	: PRIOR APPLICATION NUMBER: 60/258,928	
	: PRIOR FILING DATE: 2000-12-29	
	: PRIOR APPLICATION NUMBER: 60/259,415	
	: PRIOR FILING DATE: 2001-01-02	
	: PRIOR APPLICATION NUMBER: 60/259,785	
	: PRIOR FILING DATE: 2001-01-04	
	: PRIOR APPLICATION NUMBER: 60/269,814	
	: PRIOR FILING DATE: 2001-02-20	
	: PRIOR APPLICATION NUMBER: 60/279,832	
	: PRIOR FILING DATE: 2001-03-29	
	: PRIOR APPLICATION NUMBER: 60/279,833	
	: PRIOR FILING DATE: 2001-03-29	
	: PRIOR APPLICATION NUMBER: 60/279,863	
	: PRIOR FILING DATE: 2001-03-29	
	: PRIOR APPLICATION NUMBER: 60/283,889	
	: PRIOR FILING DATE: 2001-04-13	
	: PRIOR APPLICATION NUMBER: 60/284,447	
	: PRIOR FILING DATE: 2001-04-18	
	: PRIOR APPLICATION NUMBER: 60/286,683	
	: PRIOR FILING DATE: 2001-04-25	
	: Remaining Prior Application data removed - See File Wrapper or PALM.	
	: NUMBER OF SEQ ID NOS: 411	
	: SOFTWARE: PatentIn Ver. 2.1	
	: SEQ ID NO 41	
	: LENGTH: 8487	
	: TYPE: DNA	
	: ORGANISM: Homo sapiens	
	US-10-038-854-41	
	Query Match 40.7%; Score 3398.8; DB 17; Length 8487;	
	Best Local Similarity 67.0%; Pred. No. 0;	
	Matches 5008; Conservative 0; Mismatches 2362; Indels: 102; Gaps 9;	
QY	880 CGGGCACTTCCTCCTTCAAAGCCCTGG--AGGCACCTCCCCGCTCTTTCTGCCACCACATCACCC	936
Dd	766 CAGGCAATTCTTAATCAAACAAGGAACAGGTACAAAGCCACTGTTCACTAGTACTGCAACCCC	825
QY	937 AGGTTACCACTGACGTCCAGCACAGTGACTCTTCTCCCGCCCGACCCCTGCCCCCGCAG	996

937 AGGGTACCCCACTGACGTCGAGACAGTGTA

QY 2077 TGTGGATGGGGAGGACCAACTGCGAGACCCCGAGGCGCACATGCTTAGACAGTGTTC 2136
 Db 1918 TCCAGGATGGGGAGGTAGCAATTTGAAATACCTGAAGACCATGTGTCAGACACAGTGTTC 1977
 QY 2137 AGGCCACGGAACTTCTCTCCGGHACACGGGCTTTGAGCTGTGACCCAAAGCTTGGACTGG 2196
 Db 1978 CGGCCACGGAACTGATCTTCAAGAAAGTGGCTCTGTCAGCTGTGAACCTTAACCTGGACTGG 2037
 QY 2197 ACAGCACTGTTCTATCAGATCTGTGCTGCCAGTGTGGTGGCCATGGCGGTGTGCGTAGG 2256
 Db 2038 CCCAGACTGCTCAACGAATATGTTCTGTGGACTGTGGCTCACAGGCGTGTGCAATGG 2097
 QY 2257 GGGCACTGCCGCTGCGAGAGTGGCTGGATGGGGGAGCTGCGACACAGCGGCGCTGCCA 2316
 Db 2098 GGGGACGTGTGCTGTGAAGAGGCTGGACGGGCGCAACCTGTAATCAGAGAGCTGCCA 2157
 QY 2317 CCCGCGCTGTGCCAGCATGGGACCTGCCGCGAGCGCAAGTGGAGTGGAGCGCCCTGGCTG 2376
 Db 2158 CCCCCTGTGCCGAGCAGCGGACCTCAAGGATGGCAAGTGGATGCGACGCAATGGCTG 2217
 QY 2377 GAATGGCGAACTGCAACCATCGCTCACTATCTGGATAGGCTAGTTAAAGAGGGTTGCC 2436
 Db 2218 GAATGGAGACACTGCCTATC-----GAGGGTTGTCC 2250
 QY 2437 TGGGTTGTGCAATGGCAACGGCAGATGTACTTAGACCTGAATGGTTGGCACTGCTCTG 2496
 Db 2251 TGGTCTGTGCAACAGCAATGGAAATGTACCTGGACCAAAATGGCTGGCAATTTGTGTG 2310
 QY 2497 CCAGCTGGCTGGAGAGGAGTGGCTGTGACACTTCCATGGAGACTGCTTCCGGTGCACAG 2556
 Db 2311 CCAGCTGGATGGAGAGGAGCGCTGTGACGTAGCATCGAGACTCTTTGACACATAG 2370
 QY 2557 CAAGACAATGATGGAGATGGCTGTGGACTGATGGACCCCTGACTGCTGCCCTCAGCC 2616
 Db 2371 CAAGGACAATGAAGAGATGGACTCATTGACTGCGATGGATCCCGATTTGCTGCCCTACAGAG 2430
 QY 2617 CCTGTGCCATATCAACCGCTGTGCTGTGCTCCCTAACCTCTGGACATCATCCAGGA 2676
 Db 2431 TTCTGTGCAAGATCAGCCCTATTGTGCGGAGCTGCCGGATCTTCAGGACATCATTAGCCA 2490
 QY 2677 GACACAGTCCCTGTGTACAGAGAACCTAACCTCTTCTATGACCGCATCAAGTTCTCT 2736
 Db 2491 AAGCCTTCAATCGCCTTCTCAGCAAGCTGCCAAATCCTTTATGATCGAATCAGTTCTCT 2550
 QY 2737 CGTGGGAGGACAGCAGCACATAATCCCGGGGAGAACCCCTTTGATGGAGGCTATGC 2796
 Db 2551 TATAGGATCTGATAGCACCATGTTATACCTGGAGAAAGTCCCTTCAATAAGAGCCTTGC 2610
 QY 2797 TTGTGTTATTCTGTGGCCCAAGTGTGACATCAGATGGAAACCCCTCGTGGTGTGTAACAT 2856
 Db 2611 ATCTGTATCAGAGGCCAAGTACTGACTGTGTGATGGAATCCACTATTGGAGTAAATGT 2670
 QY 2857 CAGTTTTGTCAATACCTCTCTTTGGATATACATCAGCAGCAAGATGGAGCTTTGA 2916
 Db 2671 CTCGTTTTTCATTTACCCAGAAATGGATATATATTACCCCGCAGGACGGAATGTTGA 2730
 QY 2917 CTTGGTGACAAATGGCGGCATCTCCATCATCTGCGGTTCCGAGCGGCACCTTTTCATCAC 2976
 Db 2731 CTTGGTGGCAATGGTGGGCCCTCTCTAACTTTGGTATTGTAACGATCCCATTTCTCTAC 2790
 QY 2977 ACAGGACACACCTGTGTGGTGGCATGGGATCGCTTTTGTGTCATGGAAACCATCATCAT 3036
 Db 2791 TCAGTATCATACTGTGTGGATTCATGGAATGTCTTTATGTGATGGATACCCCTAGTCAT 2850
 QY 3037 GAGACATGAGGAGAAATCAGATTTCCAGCTGTGACCTGTGAGCAATTTTGGCCGCCCAACCC 3096
 Db 2851 GGAGAAAGAGAAATGCAATTTCCAGCTGTGATCTGAGTGGATTTGTGAGGCGCAATCC 2910
 QY 3097 AGTCGTCTCTCCATCCCATGAGCTCTTTGGCAGCTCTGTGCGAGAAAGGCCCAT 3156
 Db 2911 CATCATGTGTCTATCATCTTATCCACCTTTTTCAGATCTTCTCCTCTGAAGACAGTCCCAT 2970

QY 3157 TGTGCCGAAATTCAGGCTTTGACGAGGAAATCTTATCTCTGGCTGCAAGATGAGGCT 3216
 Db 2971 CATTTCCGAAACACAGGTACTTCCAGGAGAACTTACAAATTCAGGAAACAGATTTTGAACCT 3030
 QY 3217 GAGCTACCTGAGCAGCGGACCCCTGGCTACAAATCTGTCTGCTGAGGATCAGCTTCAACCA 3276
 Db 3031 CTCCTACTTGTAGTTCAGAGACTGAGGATATAGTCACTTCTCAAGATCACCATGACCCA 3090
 QY 3277 CCCGACCATCCCTTCAACCTCATGAAGTGCACCTCATGCTAGCGTGGAGGCGCCCT 3336
 Db 3091 GTCATATTATTCATTTAATTAATGAAGTTTCATCTTATGCTAGCTAGTAGGAAGACT 3150
 QY 3337 CTTGAGGAAGTGGTTCGCTGAGCGCCGAGACCTGCTCTATTTATTTATTTGGAACAGAC 3396
 Db 3151 CTTCCAAAGTGGTTCCTGCTCACCACCTTGGCTTATCTTTCATATGGGATAAAC 3210
 QY 3397 AGACCTCTACCAACAGAGGTGTTTGGGCTTTTCAAGAGCCTTTGTTTCCGTGGGTTATGA 3456
 Db 3211 AGATGCATATATCAGAAAGTCTATGGTCTATCTGAAGCTGTTGTCAGTTGGATATGA 3270
 QY 3457 ATATGAATCCTGCCAGATCTAATCTGTGGGAAAGAAAGAACACAGTGTGCGAGGCTA 3516
 Db 3271 GTATGAGTCTGTTTGGACCTGACTCTGTGGGAAAGAGGACTGCCATTTCTGAGGCTA 3330
 QY 3517 TGAATTCGACCGCTCCAAAGCTTGGAGGATGGAGCTAGACAAACATCATGCGCTCAACAT 3576
 Db 3331 TGAATGATGGTTCACACATGGGTGGCTGGACATTAGATAAATCATCAGTGTGGATGT 3390
 QY 3577 TCAAGTGGTATCTCTGCACAAAGGGAATGGGAGAAACAGTTTGTCTCAGCAGCTCC 3636
 Db 3391 ACAGAACGCTTACTGTACAGGGAACCGGGGAAACAGTTTCTCTCCAGCAGCTCC 3450
 QY 3637 TGTCAATGGGAGCATCATGGGCAATGGCCCGGAGAGAGCATCTCTGCCCGCTCAAC 3696
 Db 3451 AGTGTGATGATGATCATGGGCAATGGGCAAGGCGGAGCATTTCTTCCCGCTGCAAC 3510
 QY 3697 CGGCTTCTGACGCGCAACAGCTCTTGGCCCGCAGTGGCCCTCACCTGTGTGCTGACGG 3756
 Db 3511 TGGTCAAGCTGATGGTAACAGTTTACTGGCCCGCTAGCTTGTGGATCGATGG 3570
 QY 3757 GAGCTCTATGTGGTGTATTTCAACTACATTAGAGGATCTTCCCTCTGGAATGTCAC 3816
 Db 3571 CAGTCTGTACGTAGCGGATTTCAACTACGTGCGCGGATATTCCTTCTGGAAATGTAAC 3630
 QY 3817 CAACATCTCTAGCTGAGGAATAAGATTTTACAGACTAGTTCACAGTCCAGCACAAATA 3876
 Db 3631 AAGTGTCTTAGACTAGAAATTAAGATTTTAGACATAGCAGCAACCCAGCTCATAGATA 3690
 QY 3877 CTACTGCGCACAGACCCCATGAGTGGGCGCTTCTCTTTCTGACAGCAACAGCCGCGG 3936
 Db 3691 CTACCTTCAACCGGATCCAGTCCGCGGAGATCTGAGCTTTCTGACACAAACACCCGAG 3750
 QY 3937 GGTCTTTAAATCAAGTCCACTGTGTGTGAAGAGCTTGTCAAGAACTCTGAGGTGGT 3996
 Db 3751 AATTATGCGCCAAAGTCACTTACGGGGGCAAGACTTGTGACTTAAATGACAGAGTCTGT 3810
 QY 3997 TCGGGGACAGGTGACAGTCCCTCTCTTTGATGACACTCGCTGCGGGGATGTTGGGAA 4056
 Db 3811 CGAGGACAGGGGAGCAATGCCCTTCCGTTGACGAGCGGAGATGTGGGATGGAGGAA 3870
 QY 4057 GGCACAGAGGCACTCTACCAATCCAGGGGTATTACAGTGCAGCAAGTTTGGGCTGAT 4116
 Db 3871 GCGGTGGAAGCCACACTCATGAGTCCCAAGGAATGGAGTTGATGAAGATGGAATTAAT 3930
 QY 4117 CTACTTCGTGATGGCAGCCATGATCAGACGATCGATCAGAAATGGGATCATCTCCACCT 4176
 Db 3931 CTACTTTGTTGATGGAACCATGATTAGGAAGTTGACCAAAATGGAATCATATCAACTCT 3990
 QY 4177 GCTGCGCTCTAATGATCTCATCAGCCCGGCGCACTCAGCTGTGATTTCTGTCTCATGATAT 4236
 Db 3991 TCTGGCTCTAACGATTTGACTTCCAGCCAGACCTTTAACTTTGTGACACAGCATGCAAT 4050
 QY 4237 TTCCAGGTAAGATGGAGTGGCCCGCCAGACACTTATGACCATCAACCAATGGACAACTCACT 4296

	Db	4051	CAGCCAGGTCACGTCTGGAAATGGCCCACTGACCTTAGCCATTAAACCCTATGGATAACTCCCAT	4110
	Qy	4297	TTATGTCTCGACAACAATGTGTGCTGTGCAAATCTCTGAAAACCAACAGAGTGGCATTGT	4356
	Db	4111	TTATGTCTCGAATAATATGTAGTTTTACAGATCACTGAANAATCGTCAAGTTCGATTCG	4170
	Qy	4357	CGCGGAGGCCAATGACACTGCAGGTCCCTGGATGTGACCACTTCTCTGTAAAGCAAGGT	4416
	Db	4171	TGCTGGACGGCCCAATGCACCTGTGAGTTCCGGAGTGG---AATATCTCTGTGGGGAAGCA	4227
	Qy	4417	GGCCATCCACGCAACCTGGAGTCAGCACCGCTTTGGCTGTGTTTACACAATCGGGTCCCT	4476
	Db	4228	CGCGGTGCAGACAACACTGGAAATCAGCCAATGCCATTCGTGTCTCTACAGTGGSGTCTCT	4287
	Qy	4477	GTAATATGTGAGACTGATGAGAAAAAGATCAACCCGATCAGCAGGTCAACACTAGTGG	4536
	Db	4288	GTAATTTACTGAAACTGATGAGAAGAAAAATTAACCGGATAAGGCAGGTCAACACAGATGG	4347
	Qy	4537	AGAGATCTCACTCGTTGCTGGGGCCCCAGTGCTGTGACTGTATAAATATGATGCCAACCTG	4596
	Db	4348	AGAAATCTCTTTAGTGGCCGGAATACCTTCAGAGTGTGACTGCAAAAATGATGCCAACCTG	4407
	Qy	4597	TGATGTGTTTTCTGGAGACGATGGTTATGCCAAGGATGCAAGGTTAAATACCCCATCTTTC	4656
	Db	4408	TGACTGTTACAGAGTGGAGATGGCTACGCCAAGGATGCCAAACTCAGTSCCCCACTCCTC	4467
	Qy	4657	CTTGGCTGTGTGCTGATGGGAGACTCTACGTGGCCGACCTTGGGNAACATCCGATTCG	4716
	Db	4468	CTTGGCTGTCTTCCAGATGGTACACTGTATATTGACAGATCTAGGGAATATCCGATPCG	4527
	Qy	4717	GTTTATCCGGAAGAAACAAGCCTTTCTTCAACCCAGAAACATGTATGAGCTGTCTTCACC	4776
	Db	4528	GGCTGTGTCAAAGAATAAGCCTTTACTTAATCTATGAACCTTCTATGAAGTTCGGTCTCC	4587
	Qy	4777	AATTTGACAGGAGCTCTATCTGTTTGATACACCGCGGAAGACCTGTACACCCAAAGCCT	4836
	Db	4588	AACTGATCAAGAACTCTACACTCTTTGACATCAATGTGTACTCACCAAATATACTGTAACTT	4647
	Qy	4837	GCCACACGAGAGACTACCTGTACAACTTCACCTACCTGGGAGCGGACATCACACTCAT	4896
	Db	4648	AGTCACTGGTGATTAACCTTTACAATTTAGCTACAGCAATGACAATGATTAATCTGCTGT	4707
	Qy	4897	CACAGACAACAATGGCAACATGTTAATGTCCGCCGAGACTCTACTGGGATGCCCTCTG	4956
	Db	4708	GACAGACAGCAATGGCAACACCCCTTAGAATTAAGACGGGAGCCCAAAATCGCATGCGAGTCCG	4767
	Qy	4957	GCTTGGTGTCCAGATGGCCAGGTGTACTGGGTGCAACATGGGACCAACACAGTGCATCAA	5016
	Db	4768	AGTGGTGTCTCTGATTAACCAAGTGTATGTTGTGACAAATAGGAACAAATGGATGTTTGA	4827
	Qy	5017	GAGTGTGACCACAACAAGGACACGAGTTGGCCATGTATGACATACCATTGGCAATTCGGCCCT	5076
	Db	4828	AGGCATGACTGCTCAAGGACTGGAAATTAGTTTGTGTTTACTTACCATGGCAATAGTGGCCT	4887
	Qy	5077	TCTGGCAACCAAAAGCAATGAAAACGATGGACACAATTTTATGAGTACACAGACTTTGG	5136
	Db	4888	TTTAGCCACTTAAAGTGTATGAAACTGTGATGGACAACGTTTTTTTGACTATGACAGTGAAGG	4947
	Qy	5137	CCGCCTGACAAATGTGACCTTCCCTACTGCGCCAGGTGAGCAGTTTCCGAAAGTCATACAGA	5196
	Db	4948	TGCTCTGACAAATGTTACGTTTCCAACTGGAGTGGTCACAAACCTGCATGGGACATGGA	5007
	Qy	5197	CAGTTTCACTGATGTCAGGTAGAGACTCCAGG---AAGATGATGTACCAATTAACCAAC	5253
	Db	5008	CAAGGCTATCACAGTGGACATTCAGTTCATCTAGCCGAGAAAGATGTCACGATCACTTC	5067
	Qy	5254	CAACCTGTCTGCCCTCAGCGCCTTCTACACACTGTCTCAAGACCAAGTCCGGAACAGCTA	5313
	Db	5068	AAATCTGTCTCGATCGATTTCTTTACACCATGGTTCAGATTCAGTTTAGAAACAGCTA	5127
	Qy	5314	CTACATCGGGCCGATGCGCTCTCTTTCGGCTGTCTGCGCCAAACGGCATGAGGTGGCGCT	5373

Db	5128	CCAGATTGGTTATGACGGCTCCCTCAGAAATTACTATAGCCAGTGGCCTCGACTCACACTA	5187
Qy	5374	GCAGACTGAGCCCCACTTGTGTGCTGGCACCGGTCAACCCCAACCGTGGGCAAGAGGAATGT	5433
Db	5188	CAAAACAGAGCCGACGTTCTGCTGCGCACCGCTAATCCGACGGTTTGCCAAAAGAAACAT	5247
Qy	5434	CACGCTGCCATCGACAAACGGCTCAACCTGTGGAGTGGCGCCAGCGCAAAAGACAGGC	5493
Db	5248	GACTTTGCTGGCGGAACGGTCAAAACATTGGTGGAAATGGAGATTCGAAAAAGACAAGC	5307
Qy	5494	TCGGGGCCAGGTCACTGTCTTTGGGCGCGGCTGCGGGTGCACAAACGGAATCTCCTATC	5553
Db	5308	CCAAGGAAAGTCAATGTCTTTGGCGCAAGCTCAGGGTTAATGGCAGAAAACCTCTTTC	5367
Qy	5554	TCTGGACTTGTATCGGTTACACGACAGAGAAAGATCTATGATGACCAACGCAAGTTCAAC	5613
Db	5368	AGTTGACTTTGATCGAACAACAAGACAGAAAAGATCTATGACGACCAACGGTAAATTTCT	5427
Qy	5614	CCTTCGGATTCTGTACGACCGAGCGGGCGGCCAGCCTCTGTGTCAACGACGACAGGCT	5673
Db	5428	ACTGAGGATCGCCTACGACACGTCGCGCACCGACTCTCTGGCTGCCAAGCAGCAAGCT	5487
Qy	5674	GAATGGTCAACGTTGACATPACTCCCTCGTGGGGTTACATTCTGTGGGATCCAGAGGGGAT	5733
Db	5488	GATGGCCGTCAATGTCACTTATTCACAGTCAATTCACAGTCAAAATGGCCAGCATCCAGCGAGCAC	5547
Qy	5734	CATGCTCTAAGAAATGGAATACGACACGAGCGCGCCGATCACATCCAGGATCTTCCTGA	5793
Db	5548	CATAGCGGAAGTAGATTATAGCGCAGGGAGATCGTGTCTCTGGGTCTTTGCTGA	5607
Qy	5794	TGGGAAGACATGGAGCTACACATCTTAGAGAAAGTCCATGGTGCTCTACTACACAGCCA	5853
Db	5608	TGGTAAACATGAGTTACACATATTAGAAAAGTCCATGGTTCTTCTGCTTCATAGCCA	5667
Qy	5854	GAGCGAGTATCTTTGAGTTTCACAGAATGACCGCTCTCTTCTGTGACGATGCCAA	5913
Db	5668	CGCGCAGTACATCTTGAATACCATATGTGGGACCGCTGTCTGCCATCACCAATGCCAG	5727
Qy	5914	CGTGGCGCGCAGACACTAGAGACCATCCGCTCAGTGGGTACTACAGAAACATCTATCA	5973
Db	5728	TGTGGCTCGCCACACCATGCAGACCATCCGATCCATTGGCTACTACCGCAACATACAA	5787
Qy	5974	GCCCTTGAGGGCAATGCCCTCAGTCAACAGGACTTCACTGAGGATGGGCACCTCTTCA	6033
Db	5788	CCCCCGGAAGCAACGCCCTCCATCATCGGACTACACGAGGAAGGGCTGTTCTACA	5847
Qy	6034	CACCTTCTACTGGGCATCGCCGACGGGTGATATACAAAGTATGGCAAACTGTCAAAGCT	6093
Db	5848	AACAGCTTCTTGGGTACAAGTCGGAGGTCTTATTCAAATACAGAGGCAGACTAGGCT	5907
Qy	6094	GGCAGAGACGCTCTATGACACCAACCAAGTCACTTCACTTATGACAGACGGCAGCAT	6153
Db	5908	CTCAGAAATTTTATATGATAGCACAGAGTCAGTTTACCTATGATGAACACAGCAGAGT	5967
Qy	6154	GCTGAACACCATCAACCTACAGAAATGAGGCTTCACTGCACCATCCGCTACCGTCAGAT	6213
Db	5968	CCTTAAGACAGTAAACCTCCAGAGTGAATGGTTTTATTTCACCATTAGATACAGGCAT	6027
Qy	6214	TGGGCCCTGATTGACCGACAGATCTTCGCTTCACTGAGGAAGGATGGTCAACGCCCG	6273
Db	6028	TGTTCCCTGATTGACAGGCAGATTTTCGCTTTAGTGAAGATGGGATGTTAATGCAAG	6087
Qy	6274	TTTTGACTACAACTATGACAAACAGCTTCGGGTGACCGATCGACGCTGTGATCAACGA	6333
Db	6088	ATTTGACTATAGCTATGACAAACAGCTTTCGAGTGAACAGCATCGACGGTGTGATCAATGA	6147
Qy	6334	GACCCCACTGCCATTGATCTCTATCGCTATGATGTGTGAGGAAGACAGACAGAGT	6393
Db	6148	AACGCCACTGCCATTGATCTGTATCAGTTTGTATGATCTTCTGGCAAAAGTTGAGCAGTT	6207
Qy	6394	TGGGAAGTTTGGTGTCACTTACTATGACATTACCAAGATCATCACACAGCTGTCAATGAC	6453
Db	6208	TGGAAGTTTGGAGTTATATATATGATATTAACCAAGATCATTTTACAGCTTAATGAC	6267

Qy	6454	CCACACCAAGCAATTTTGATGCATATGCGAGATGAAGAAAGTCAGATGAGATCTTCCG	6513
Db	6268	CTATACCAAGCACCTTTGATGCTCATGCGCGGTATCAAGGAGATTCAATATGAGATATTACG	6327
Qy	6514	CTCGCTCATGTACTGGATGACCGTCCAGTATGATAACATGGGCGAGTACTGAAGAAGGA	6573
Db	6328	GTGCTCATGTACTGGATTCAATTCAGTATGATPAACATGGTCCGGTACCAAGAGAGA	6387
Qy	6574	GCTGAAGGTAGGACCTTACGCCAATACCCTCGCTACTCTCTATGAGTATGATGCTGACGG	6633
Db	6388	GATTTAAANATAGGGCCCTTTGCCAACACCAATAATGCTTTATGAAATATGATGTTGATGG	6447
Qy	6634	CCAGCTCGACAGAGCTCTCCATCAATGACAAGCCACTCTGGCGGTACAGCTACGACCTCAA	6693
Db	6448	ACAGCTCCAACAGTTTACCTCAATGAAAGATATATGTGGCGGTACAATACCATCTGAA	6507
Qy	6694	TGGAAACCTGCATTTACTGAGCCCTGGGAACAGTGCACGGCTCACACCCTACCGTATGA	6753
Db	6508	TGGAAACCTCCATTTACTGAACCAAGTAACAGTGCAGCTGACACCCCTTCGCTATGA	6567
Qy	6754	CATCCGGACCGCATCACTCGGTGGGTGACGTGCANATACAAGATGATGATGAGATGCGTT	6813
Db	6568	CCTGCGAGACAGAAATCACTGCATGGGTGATGTTCAATATCGGTTGGATGAAATGGTTT	6827
Qy	6814	CCTGAGCAGCGGGCGGTGATATCTTTTGAGTACAACTCAGCTGGCGTGTCTATCAAGGC	6873
Db	6628	CCTACGTCAAAGGGGCACGGAATCTTTGAATATAGCTCCAAGGGGCTTCTAACTCGAGT	6687
Qy	6874	CTACAAACGGGCTGGCAGCTGGAGTGTCAAGTACCGCTACATGAGTGGCGGGCGCGCT	6933
Db	6688	TTACAGTAAAGGCGAGTGGCTGGACAGTATCTACCGTTATGACGGCGCTGGAAAGGCGGT	6747
Qy	6934	GTCCAGCAAGCAGCAGCCACGACCACTTCGAGTCTCTTATGCGACACTGACCAACCC	6993
Db	6748	TTCTAGCAAAAACAGTCTAGACAGCACCTTGCAGTTTTTTTATGCTGACTTAACCTATCC	6807
Qy	6994	CACCAAGTCAACCACTGTACAACCACTCCAGCTCTGAGATCACCTCCCTCTACTACGA	7053
Db	6808	CATTAGATTACTCATGCTGTACAACCAATTGAGTTCGAAATTAACCTCCCTGATTATGA	6867
Qy	7054	CTTGCAAGGACACCTCTTTGCCATGGAGCTGAGCAGTGGTATGAGTTTTACATAGCTTG	7113
Db	6868	TCCTCCAGGACATCTTTTGCATGGAATCAGCAGTGGGATGAAATCTATATTGCATC	6927
Qy	7114	TGACAAATCGGACCCCTCTTGCTGTCTTTAGTGGAAACAGTTTGTATGATCAAGCAAT	7173
Db	6928	GGATAACACAGGACACCACTGGCTGTGTTCAGTAGCAATGGGCTTATGCTGAAACAGAT	6987
Qy	7174	CCTGTACACAGCCTATGGGAGATCTACATGGATACCAACCCAACTTTTCAGATCATCAT	7233
Db	6988	TCAGTACATGTCATATGGGAAATCTATTTTGACTCTAATATTGACTTTCAACTGGTAA	7047
Qy	7234	AGGTCACATGGTGGCCCTATATGATCAACATCCAGCTTGTCCCATGGCGCGGAGGA	7293
Db	7048	TGGATTTATGSGTGGCCGTATGACCCACATCACCAAAATTAATCCACTTTTGGAGAAAGA	7107
Qy	7294	TTATGATGTCTGGCGGACGCTGGACTAGCCAGACACGAGCTGTGGAAAGCACCTTAG	7353
Db	7108	TTATGACATTTTGGCAGGACGTTGGACCAACCTGACATAGAATCTGGA--AAAGAT	7164
Qy	7354	TAGCAGCAACGTGATGCTTTAATCTCTATATGTTTCAAAAAACAACCCCATCAGCAA	7413
Db	7165	TGGGAAGGACCCAGCTCCTTTTAACTTGTCATGTTTAGGAATAACAACCCCTCAAGCAA	7224
Qy	7414	CTCCCAGGACATCAAGTGTCTATGACAGATGTTAACAGCTGGCTGCTCACTTTGGATT	7473
Db	7225	AATCCATGACGTGNAAGATTACATCAGAGATTTACAGCTGGCTGGTGCATTTGGTTT	7284
Qy	7474	CCAGCTACAAACGTGATCCCTGGTTATCCAAACACAGACATGGATGCCATGAAACCTC	7533
Db	7285	CCATCTGCAAAATGCTATTCCTGGATTCCCTGTTCCCAATTTGATTTACAGAACCTTC	7344

```

RESULT 9
US-09-808-602-78
; Sequence 78, Application US/09808602
; Patent No. US202020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkes, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20202015511
; FILE REFERENCE: 15965-697 CIP
; CURRENT APPLICATION NUMBER: US/09/80

```

; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 78
; LENGTH: 8689
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-09-808-602-78

Query Match 40.78; Score 3397.8; DB 9; Length 8689;
Best Local Similarity 64.8; Pred. No. 0; Mismatches 2712; Indels 254; Gaps 19;
Matches 5469; Conservative 0

QY 35 ATGACGTGAAGGAGAGAGACCTTACCGCTCGCTGACCCGCGCGCGGACGCGGAGCGC 94
Db 133 ATGATGTGAAGATCGCGACATCGCTCTTTGACACGGGACGGTGTGCGAAGAGTGT 192

QY 95 CGCTACACAGCTCGTCCGGGACAGGAGAGGGGAAAG---CCCGCAGAAATCGTAC 151
Db 193 CGCTACACAGCTCGTCTCTGGACAGTGAAGGACTGCGGTGCGCCACGCGAAGTCTCTAC 252

QY 152 AGCTCCAGCGAGACCCCTGAAGGCTACGACGAGCGCGCTAGCCTATGCGACCGC 211
Db 253 AGTTCAGTGAGACCTTGAGGCTTATGACATGACGAGATGACTATGAAACCGA 312

QY 212 GTCAGGACATTTGCGCGAGGAGCGAGGAATTTGCGCGACAGGTGCCAATTCACC 271
Db 313 GTCAGACCTGTGTCACCGGAGTCCGATGAGTCTTTAGACAGGGGCTAATTTACC 372

QY 272 CTGCGGAGCTGGGCTGGAGAGTAACGCCCCCTCAGGGACCTGTACCGGACAGAC 331
Db 373 CTGCGAGATTTGGAAATCTCGAG---CCCTCCCGACACCGAAGTGTACTGTTCCGAC 429

QY 332 ATTGGCTGCCCAATCGGCTACTCCATGGGGCTGGCTCTCATGCGGACATGAGGCT 391
Db 430 ATGGGATCTCCACAGGCTACTCCCTGAGCACTGGGTCTGATGCGGACTCGACACC 489

QY 392 GACACGCTGTCTCCCTGAGACCCCGTGCCTGTGTGGGCGCGAGACACGGTCAGG 451
Db 490 GAGGGAGGATGTCTCCAGAACATGCCATCAGACTGTGGGACGAGGGATAAATCGAG 549

QY 452 CGCAGCTCTGCTGTCCAGCGGGCCAAATTCAAATCTCAGCTCAGCAGACCGGAGCAT 511
Db 550 CGCAGCTCTGCTGTCCAGCGGAGACTGAGCCCTTACTCTGACTGATTTGACAT 609

QY 512 GAAAAAAGTGAAGTATCCGGGCGCC---TGCAAGACACCGCGGCTCCGGAC 567
Db 610 GAAAAATAATCGGATGACGACAAATGCTGACCCCAATTCACCTACATCTCCTGTCGCTC 669

QY 568 GCCGCGCGCGCTCTCGACCGCCACACCCCAACCGACGACCGCGGCTCCATTAA 627
Db 670 CTCCTATCTGCTAGCTGCTAGCTCTCCATTAATCTCCAGCTTGTAGTGC---AGATGCC 728

QY 628 CTCCTGAAACGGGGCAATCTACGCGGAGGAGCAACCCCGAGCGG--- 674
Db 729 ATTGCTAGACGACAAACCTCCCATCAGATCATGAGACACCAACCCCGATGAGGAATCTC 788

QY 675 -----CCCCACGGACCACTCGCTCTCGGAGAGCCCTTGGCGGGCGGCC 722
Db 789 CCTAATTCATCTGCTCAGACATGCTCAGGGCCCCAGCAAGCCTCCAGTAGTGGGCC 848

QY 723 AGGAGCTGCGCCACGGGAGAAATGGGCTGTCAACAGCAACATCCCGCTGGAGACCA 782
Db 849 TCCGACACCAACAGCCAGTCAAGCTGAGGCCCTCTGCGACCTCTCTCATAAACACAC 908

QY 783 GGAACCTAGGACAGGAGCAATTCATAGGACATTTGAGGACACCTCATTTGAGATGACA 842
Db 909 CTGTGCCACCACTCTCTGCTGCCAACTCCCTCAACAGAAATCTACTGACCAATCGGG 968

QY 843 TTCTCGGGCGCTCCCGCCA----- 861
Db 969 GAGTCAAAATCCACGCCCGCAGCTCTGACCCCAATGACCTGGCCACCAACGCGGAGTCCGT 1028

QY 862 -----TGATGGGCTTACAGTGACGGCACTTCTCT 891
Db 1029 TCAGCTCCAGGACAGCTGGGTGCTGAAACAGTAACGCTGCCGTGGAGACGGGCACTTCTCT 1088

QY 892 CTTCAA---GCCTGGAGGACACCTCCCGCTCTTTGACACCAATCACCAGGTTACCACT 948
Db 1089 CTTCAAGAGCTCTCCGGAAGCACACCCCTGTTACGACAGCTCTTCTCCAGGATACCCCTT 1148

QY 949 GAGCTCCAGCAGAGTACTCTCTCCGCGCGGACCCCTGCGCCGACGACGACCTTCCCGG 1008
Db 1149 GACCTCAGGAGCCGTTTATACACCAACCCCGCTGCTGCCACGGAATACATTTCTCTAG 1208

QY 1009 GCGCGCTTTAACTCAAGAAACCCCTCCAAGTACTGTAACTGGAAGTGGCAGACCCCTGAG 1068
Db 1209 GAAGGCTTCAAGCTGAAGAAACCCCTCCAATCTGCACTTGGAAATGCGCGCCCTGTC 1268

QY 1069 CCGCATTCGTATCTCAGCCACTCTGGTCACTCTCTGCTGGCATACTTTGTCGCACTGCACT 1128
Db 1269 TGCCATTCGCGTGCCTCTCTTGGCCATTTTGTGCGCTATTTCTAGCAATGCACTCT 1328

QY 1129 GTTTCGCTTAACTGGCACCTTGACCGATGGAGGGGAGATGTATGAGATCACGAGGA 1188
Db 1329 GCTCGCATCTAATTTGCAACTCCAGCCGCGAGATGGACACACCTTTAAACA-----ATGG 1382

QY 1189 CACAGCCAGCACTGGCTGTGCCAACCGACCTCTCCCTATATACCCCTCAGGGGCACTGG 1248
Db 1383 CGTAAGGACCGGCTTTACCAGGAACGATGATGTGGCAACAGTGCCTCTGAGGCAAGT 1442

QY 1249 CTTAGAGACCCCTGACAGGAAAGGCAAGGAACACAGAAAGGCAAGCCAGTAGTTCTT 1308
Db 1443 GCCCTGGTCTGTGAA--- 1459

QY 1309 TCCAGAGACAGTTTTCATAGATTTCTGGAGAAATGATGTGGGAAAGCGGAGCTCCAGAA 1368
Db 1460 -----ACAGCAGCATAGACAGCGGAGGAGAGTGGTTCGACGGTGCACAGGA 1511

QY 1369 GATTCCTCTGCGACTTTCTGAGATCTCAAGTGTTCATAGACCATCTCTGTGCACTGAA 1428
Db 1512 AGTCCACACGAGGCTTTTGGAGTCCAGATTCACATCAGTCACTGAGCTCTAGTTCTTAA 1571

QY 1429 ATTCAATGTCTCTGGGAAAGGAGCCCTGTTGGCAATTTATGSCAGAAAGGCTCC 1488
Db 1572 GTTCAACATCTCCCTGGGAGAGTGCCTCTTCCGCTCTACATAAGAGAGGACTGCC 1631

QY 1489 TCCTTCACATACAGATTTGACTTTGTGAGCTCTGATGGCAGGAGGCTCTTAACCCA 1548
Db 1632 ACCATCTCATGACAGTATGACTTCAAGAACGCTGGACGAA-----A 1676

QY 1549 GAGGCGCGGAGCTTAGAGGGAACCCCGCGCAGCTCTGGGGAACCTGTCGCCCTCCAG 1608
Db 1677 GGAGAGTGGAGTGTGTCGAGTCAACGAGGAACGCGGAGCATCCAGACCTGGTGCA 1736

QY 1609 CCATGAGACAGCTTCATCCAGTATTTGGATTCAGGAATCTGGCACTTGGGTTTTTCAA 1668
Db 1737 GAAAGAGCTGTGTTGTCGAGTACTTGGATGTGGGCTGTGGCCTCGCTCTCTACAA 1796

QY 1669 TGAAGGAGAGTCAAGTGGTTCCTTCTTCCACTGCCATGCCATGAGTCGTTGGATGATA 1728
Db 1797 TGACGGCAAGGACAGGAGTGGTCTCTTCAATACGGTGTCTTAGATTCACTGTCAGGA 1856

QY 1729 CTGCCCCAGCAACTGCTATGGCAATGGTACTGCTCTGGGAGCTCTGCCACTGCTTCTCT 1788
Db 1857 CTGTCCAGCAACTTGCACGGAACGCGGAATGGTGTCTGGACTGTGTCTACTGTTTCCC 1916

QY 1789 GGGTTTCTGGGCCCCGACTGTGGCAGAGCTCTGTCGCCCTGCTGCTGTAGCGGAAATGG 1848
Db 1917 AGGATTCCTAGTGTGAGACTCGCTTAAAGCTGCTGCCCTGTTCTGTGCACTGGGAATGG 1976

QY 1849 CCAATACATGAAGGAGAGTCTTTGTGCCAGTGGCTGGAAAGGCGCTGAGTGGAGTGT 1908

Db 1977 ACAGTATTTCCAAAGGACATGCGAGTGTCTACAGTGGCTGGAAGGACAGAAATGCGATGT 2036
 Qy 1909 GCCCACCACCAACAGTGTATCATGTGGCTGCAGCAACACCTGACCTGACATCAGGGGCAC 1968
 Db 2037 GCCCATGAACAGTGTGATCGATCTTCTGTGGGGCCAGGCTCTGTGATGATGGAA 2096
 Qy 1969 CTGCATCTGCAACCTGGCTTACAGGCGGAGAGCTGTGAGGAAGTGCAGTGTGATGACCC 2028
 Db 2097 CTGCGTGTGTGACGTGGCTTACAGGCGGAGACCTGCAAGAGAGTGGATTTGCTTGGATCC 2156
 Qy 2029 CACATGTTACAGGCGGGTGTCTGCTGTGAGAGGAGTCCATGCTTTTGTGGATGGG 2088
 Db 2157 AACCTGTCTCAGCGCATGGTGTCTGTGTGAACGAGAGTGTCTATGAGCCCCGGCTGGGG 2216
 Qy 2089 AGGCACCAACTGGGAGACCCCGGAGGCCACATGCTTTAGACCAAGTGTTCAGGCCACCGAAC 2148
 Db 2217 CGGGCTCAACTGCGAGCTGGCGAGGGTCCAGTGCACAGACCAAGTGTAGTGGCATGGCAC 2276
 Qy 2149 CTTCCTCCCGSACACCGGGCTTTGCGAGCTGTGACCCCAAGCTGACATGGAACAGACTGTTC 2208
 Db 2277 TTACCTCCCTGACTCTGGCTCTGCAACTGTGTATCCGAATTTGATGGGTCCCGACTGCTC 2336
 Qy 2209 TATCGAGATCTGTCTCCGACTGTGTGGCTGATGGCTGTGGTAGGGGACACCTGGCG 2268
 Db 2337 TGTGAAGTGTCTCAGTAGACTGTGGCACTCAGGGGCTGTGATCGGGGAGGCTGGCG 2396
 Qy 2269 CTGCGAGGATGGCTGGATGGGGGAGCTCTGCAACAGAGGGGCTGCGACCCGGCGTGTGC 2328
 Db 2397 CTGTGAAGAGGGCTGGAGAGCGCGCTTGTGACCAAGCGGTGTGCGACCCCGCTGCAT 2456
 Qy 2329 CGAGCATGGGACCTGCGCGACGCGCAAGTGCAGTGCAGCCCTGGCTGGAAATGGCGAAC 2388
 Db 2457 TGAGCAGGGACCTGTAAAGATGGCAATGTGAATCCGAGAGGCTGGAAATGGTGAACA 2516
 Qy 2389 CTGCACCATCGCTCACTATCTGGATAGGAGTGTAAAGAGGGTGGCTGGGTTGTGCAA 2448
 Db 2517 CTGCACCATTTGAT-----GGCTGCCCTGATTTGTGCAA 2549
 Qy 2449 TGGCAAGCGGAGATGTACTTAGACTGATGTTGGCACTGGCTGCGCTGCGAGCTGGGCTG 2508
 Db 2550 CGGTAAACGGAGATGACACTGGGTGAGAACAGCTGGCACTGTGTCTGCCACAGCGGCTG 2609
 Qy 2509 GAGAGGAGCTGGCTGTGACACTTCCATGGAGATGCTGCTGGGTGACAGCAAAAGACAATGA 2568
 Db 2610 GAGAGGCGCGGATGCAACGTTGCCATGGAACCTCTCGCTGTGATAACAAGGATAATGA 2669
 Qy 2569 TGGAGATGGGCTGGTGGACTGTGATGGAACCTGTGCTGCTCCAGCCCTGTGCCATAT 2628
 Db 2670 GGGAGATGGGCTGGTGGACTGCTGGAACCTGTGCTGCTCCAGTCAAGCTGTGAGAA 2729
 Qy 2629 CAACCGCTGTGCTTGGCTCCCTAACCTCTGGACATCATCCAGGAGACACAGGTCCC 2688
 Db 2730 CAGCCTGCTGTGCTGGGGTCTCGGACCCCTTGGACATCATTCAGCAAGGCCACAGACA 2789
 Qy 2689 TGTGTACAGCAGAACCTACACTCTTCTATGACCGCATCAAGTTCCTCGTGGGCAGGA 2748
 Db 2790 CTGGCTCGGGTGA-----AGTCTCTATGATGCTATCAAGCTCTTGGCAGGCAAGGA 2843
 Qy 2749 CAGCAGCACAATAATCCCGGGAGACCCCTTTGTATGGAGGCACTGTGTTATTTCG 2808
 Db 2844 CAGCAGCACAATCATCTTCTGGAGACACCCCTTCAATAGCAGCTGGTGTCTGTATCCG 2903
 Qy 2809 TGGCCAAAGTGTGATGACATCAGATGGAACCCCTCGTGTGTGGAACATCAGTTTGTCAA 2868
 Db 2904 AGGCCAAAGTAGTAACACCGATGGACCCCTCGTGGGTGTGAATGTGTCTTTGTCAA 2963
 Qy 2869 TAACCTCTCTTTGGATATACATCAGCAGGCAAGATGGCAGCTTTGACTTGTGACAAA 2928
 Db 2964 GTACCCAAAATATGGCTTACACCATCACTCGCCAGGACGGACGTTTGAACCTGATGCCAA 3023
 Qy 2929 TGGCGGATCTCCATCATCTGCGGTTGAGCGGGACCTTTTCATCACACAGGAGACAC 2988

Db 3024 TGGGGGCTTGGCTTGAATCTTCTACTTTGAGCGAGGCCCTTTCTATGAGCGGAGAGCGCAC 3083
 Qy 2989 CCTGTGGCTGCCATGGGATGCTTCTTTTGTATGGAACCATCATCATGAGACATGAGGA 3048
 Db 3084 AGTATGCCGCGTGGAAACAGCTTCTATGCAATGACACACCTGTTATGAAGACGGAGGA 3143
 Qy 3049 GAATGAGATTTCCAGCTGTGACCTGACCAATTTTGGCGGCCCCAACCCAGTCTCTCTCC 3108
 Db 3144 GAATCTCATCCCCAGCTGTGACCTGAGTGGCTTTGTCCGGCCTGATCCGATCATCATCTC 3203
 Qy 3109 ATCCCCACTGACGCTCTTCCGACAGCTCTGTGAGAGAAAGGCCCATTTGTGCGGAAAT 3168
 Db 3204 CTCTCTCTGTCCACTTCTTTCAGCGCTTCCCCTGGCGGAAACCCCATTTGTGCTGAGAC 3263
 Qy 3169 TCAGGCTTTGAGGAGGAAATCTCTATCTCTGTGCTGCAAGATGAGGCTGAGCTACCTGAG 3228
 Db 3264 CCAGGTTCTTCTATGAGGAGATCGAGCTCCCTGGCAGCAACGTGAAGCTCCGTTACCTCAG 3323
 Qy 3229 CAGCGGACCCCTGGCTACAAATCTGTCTGAGGATCAGCTCACCCACCCGACCATCCC 3288
 Db 3324 CTCAGAAACAGCAGGATCAAGTCACTGTCTGAGATCACCATGACCCAGTCTCAGGTGCC 3383
 Qy 3289 CTTCAACCTCATGAAGGTGCACTCATGTGTAGGGGTGGAGGGCGCTCTTTCAGGAAGTG 3348
 Db 3384 CTTGAACCTCATCCGGGTTCATTTGATGTTGCGTGGAGGGCATCTCTTTCAGGAAGTC 3443
 Qy 3349 GTTCGCTGACGCCACAGCTGCTCTATTTATTTTGGGAGCAAGACAGACGCTACAA 3408
 Db 3444 GTTCCAGCTTCTCCCAACCTGGCTTACATTCATCTGGGACAAAGACAGCTTATGG 3503
 Qy 3409 CCAGAAGTGTTTGGGCTTTTCAGAAAGCTTTGTTTCCGTGGGTTATGAATATGAATCTCTG 3468
 Db 3504 CCAAGGGTTATGGCTATCGGATGCTGTTGTCTGTTGGATTTTGAATATGAGACCTG 3563
 Qy 3469 CCCAGATCTAATCTGTGGGAAAAAAGAACACAGTCTCTCAGGGCTATGAATGACGC 3528
 Db 3564 CCCAGTCTCATCTGTGGGAAAAAAGACAGCCCTACTTCAAGGATTCAGGCTGGACCC 3623
 Qy 3529 GTCCAAGCTTGGAGGATGGAGCTAGACAAACATCATGCCCTCAACATTTCAAAAGTGTAT 3588
 Db 3624 TTCCAACTTGTGCTGGTCCCTGGATAGCACCAACCTCAATGTGAAGCCGAAT 3683
 Qy 3589 CTTGCAAAAGGGAATGGGAGAACCAAGTTTGTCTCTCAGAGGCTCTCTGTATTTGGGAG 3648
 Db 3684 ACTACTCAAAAGGACAGGGAGAACCAAGTTCTCTGACCCAGCAGCCGCCATCATCACAG 3743
 Qy 3649 CATCATGGCAATGGCGCGGAGAGCATCTCTCTGCCAGCTGCAACCGCTTGTCTGA 3708
 Db 3744 CATCATGGGTAACCGTCCGCGCAGAGCATCTCTCTGCCAGCTGCAATGGCCTTGTCTGA 3803
 Qy 3709 CGGCAACAAAGCTCTGCGCCCGAGTGGCCCTCACCTGTGGCTCTGACGGGAGCTCTATGT 3768
 Db 3804 AGGCAACAAACTGTTGGCCCGCTGGCTGGGATCGATGGAGCTCTTTGT 3863
 Qy 3769 GGGTGAATTTCACTAATAGAGGATCTTCCCTCTGGAATGTCAACACATCTCTAGA 3828
 Db 3864 CGGTGACTTCAATTTATTCGGCGCATCTTCCCTCTCGAAACAGTCAAGTATCTTTGGA 3923
 Qy 3829 GCTGAGGAATAAGATTTTCAGACATAGTCAAGTCCAGCAGCAAAATPACTACCTGGCCAC 3888
 Db 3924 GTTAGGAAATAAGAGTTTAAACATAGCAACAGCCAGGACACAAAGTACTTTGGCTGT 3983
 Qy 3889 AGACCCCATGAGTGGGCGGCTTCTCTCTTCTGACAGAACAGCGCGGGGCTTTTAAAT 3948
 Db 3984 GGACCTGTGATGGCTCGCTCTATGCTCTGACACCAACAGTCCGCGGATFACCGAGT 4043
 Qy 3949 CAAGTCCACTGTGTGTGTGAAGGACCTTTGTCAAGAACTCTGAGGTGTTTGGGGGACAGG 4008
 Db 4044 CAAGTCTTAAGCGGAGGCAAGACCTTGGCTGGGAATTCGGAAATTTGGCCCGGACTGG 4103
 Qy 4009 TGACAGGTGCTCCCTTTTGAATGACATCTCGTGGGGGATGTGGGGAAGGCCACAGAGC 4068
 Db 4104 CGAACAAATGTCTACCTTTGATGAAGCCCGCTGTGGGGATGGCGGGAAGGCTGTGGATGC 4163

QY	4069	CACACTACCAATCCAGGGGTATTACAGTGGAGAAAGTTTGGGCTCATCTACTTCTGGGA	4128	QY	5149	TGTGACCTTCCTTACTGTGCCAGGTGAGCAGTTCCTCGAAGTGATACAGACATTCAGTGCA	5208
Db	4164	CACCTGATAGCCCTTAGAGTATTGAGTAGAAGAAAGAGGGCTTATGATATTTTGTGA	4223	Db	5241	TGTGACTCGCCACAGGGGTGGTGACAGGCTGACCGGGAATGGAATAATCCATCAC	5300
QY	4129	TGGACCATGATCAGACGCGATCGATCAGATGGATCATCTCCACCCCTGCTGGCTCTAA	4188	QY	5209	TGTCCAGGTAGACCTCCAGCAAGGAT---GATGTCAACCAACCAACCACTGTGTGTC	5265
Db	4224	TGCCACCATGATCGGGAAGGTGACCAAAATGAAATCATCTCCACCCCTGCTGGCTCCAA	4283	Db	5301	CGTTGACATTCAGAAATCCAAACCGGTAAACGATGTCACTGTGATTACCAACCTCTCTTC	5360
QY	4189	TGATCTCACATACGCGGCACTCAGCTGTGATTTCTGTCATGATATTTCCAGGTAG	4248	QY	5266	CTCAGGCGCTTCTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGC	5325
Db	4284	TGACCTCACAGCTGTCCGACCACTGAGCTGTGACTTAGCATGAGCGTGGCCAGGTCCG	4343	Db	5361	AGTGAGGCCCTCCATACCGTGGTACAAGATCAAGTGGGAACAGCTACCAAGCTCTGAG	5420
QY	4249	ACTGGAGTGGCCACACAGACTTAGCCATCAACCAATGGCAACTCACTTTATGTCTCGA	4308	QY	5326	CGATGGCTCTTGGGGTGTCTGGCCAAACGGATGAGGTGGCGCTGCAAGCTGAGCC	5385
Db	4344	TCTAGAATGGCCGACAGACCTTCGGGTCAACCCCATGGCAATTCCTGTAGCTCTGGA	4403	Db	5421	CAAGGGACCTGCGCGTCATGTACGCCAACCGCATGGCGTCAGCTTCCACAGAGGCC	5480
QY	4309	CAACAATGGTCTCTGCAAACTCTGAAACACACAGGTGCGGATTTGTCGGGAGGCC	4368	QY	5386	CCACTTGTGCTGGCACCGTCAACCCACCGTGGCAAGGAATGTCAAGCTGCGCAT	5445
Db	4404	GAACAACGTTCATCTCGGATCAACGAGATCAACAGGTGAGCATCATCGGGAGGCC	4463	Db	5481	CAAGTCTCTCGACGCAACCTCACCCACCATCGGGCGCTGTAACTCTCCCTGCGCAT	5540
QY	4369	CATGCACTGCCAGTCCCTGGCACTGACCACTTCCTGCTAAGCAAGGTGCCATCCAGC	4428	QY	5446	CGAACACGGCTCAACCTTGTGGAGTGGCGCAGCGCAAGAGAGGCTCGGGGCCAGGT	5505
Db	4464	CATGCACTGCCAGTTCCTGGCATCGACTACT---CGCTCAGCAAGCTCGCATCCACTC	4520	Db	5541	GGAGAACGGCTGAATCCATCGATGGCGCTCAGGAAGGAACAGATTAAAGCAAGT	5600
QY	4429	AACCTGGAGTACGCCACCGCTTTGGCTGTTTCCACAATGGGGTCTGTATTTGCTGA	4488	QY	5506	CACGTCTTTGGGCGCGGCTGGGGTGCACAAACCGAAATCTCTATCTCTGGAGCTTTGA	5565
Db	4521	TGCTCTGGAGTACGCCAGCCCATCGCAATTTCTCACACCGGGGTGCTCTACATCACGA	4580	Db	5601	CACCATCTTTGGGAGGAGTTCGGGTCCACGGAAGGAACCTCTGTCCATTGATTATGA	5660
QY	4489	GACTGATGAGAAAGATCAACCGCATCAGGAGGTCAACACTAGTGGAGAGATCTCACT	4548	QY	5566	TCGGTACACGACAGAGAGATCTATGATGACACCGCAAGTTCACCTTCGGATTCT	5625
Db	4581	GACGGACGAGAAGATCAACCGCTTACCGCAGGTCAACCAACGAGAGATCTGCT	4640	Db	5661	CGAAATATCCGACTGAGAAGATCTATGACACACCGGAAGTTTCACTTGAAGTAT	5720
QY	4549	CGTTGTGGGGCCCCAGTGGCTGTGACTGTAAATATGATGCCAATGTGATTGTTTTTC	4608	QY	5626	GTACACACAGCGGGGGGCCCGCTTGTGTACCCAGCAGCAGGCTGAATGTGTCAA	5685
Db	4641	CTTAGCGGGGAGCCTCAGACTGTGACTGCAAAATGACGTCAACTGCTATCTGCTATTTC	4700	Db	5721	TTATGACAGGTGGCGCGCCCTTCTGTGCTCCACAGCAGTGGAGTGGCGGCGCTCAA	5780
QY	4609	TGAGACGATGGTATGCGAAGATGCAAGTTAAATACCCCATCTCTTGGCTGTGTG	4668	QY	5686	CGTGACATCTCCCTGGGGTTACATGTGTGGCATCCAGAGGGGCATCTGTCTGAAAG	5745
Db	4701	GGGAGATGACGATACGCCACGATGCCATCTTTGAATCTCCCGTCTCTCTTACTGTGCG	4760	Db	5781	TGTCCTCTACTTCTTCAACGGGGCGCTGGCGGGCTCCAGCGGGGCCATGAGCGAGAG	5840
QY	4669	TGCTGTGGGGAGCTCTACGTGGCGACCTTGGGAACATCCGAATTCGGTTTATCCGAA	4728	QY	5746	AATGGAATACACAGCGGGGCCCATCACATCCAGGATCTTCGCTGATGGGAAGACATG	5805
Db	4761	TCCGGATGGACCATCTACATCCAGACCTCGGGAATATCCGATCAGGCGGTGAGAA	4820	Db	5841	GACGACATTCACAAGCAAGCGCGATTGTGTCCGGAATGTTCCCGACGCGGAAGCTGT	5900
QY	4729	GAAACAGCTTTCTCTCAACCCAGAAATGTATGAGTGTCTTTCAACCAATTGACCGAA	4788	QY	5806	GAGTACACATCTTAGAAGATCTAGTGTCTGCTACTACAGCCAGAGGCTGATATAT	5865
Db	4821	AAACAAACCTGTCTTAAACGGTTCAACCAAGTATGAGGTGCTCTCCGGGAGAACAGGA	4880	Db	5901	GAGCTATTCTTACCTTGACAAAGTCCATGCTCTGCTGCAGAGCCAGCGTCTAGTACAT	5960
QY	4789	GCTCTATCTGTTGATACACCGGCAAGCAGCTGTACCCCAAGCTCCCGCAGGAGA	4848	QY	5866	CTTTGAGTTCCACAAGAAATGACCGCTCTCTCTGTGACGATGCCCAACGTGGCGCGGCA	5925
Db	4881	ACTGTAGTGTTCACGCGGATGATTCATCATGATACCGTGAGCTGTGACCGGGGA	4940	Db	5961	ATTGGAATATGATCTCTCTGACCGCTCCACGCACTACCATGCGCAGTGTGCGCCGCA	6020
QY	4849	CTACCTGTCAACTTCACTACACTGGGACGGCGACATCACACTCATCACAGACAACAA	4908	QY	5926	GACACTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCTTGAGGG	5985
Db	4941	GTAATTATACAAATTTCACTACAGCGCTGACAAATGATGACCGAGTTGATTGACAACAA	5000	Db	6021	CAGCATGTCCACGACACACTCCATTGGCTACATCCGGAACATTTACAACCCACCGGAAG	6080
QY	4909	TGGCAACATGGTAAATGTCGCGAGACTCTACTGGATGCCCTCTGGCTGTGTGTCCTC	4968	QY	5986	CAATGCCCTCAGTCATACAGGATCTTCACTGAGGATGGGCACTCTTCACTTCTTACTCT	6045
Db	5001	CGGGAATTCCTTAAAGATCCGCGGACAGCAGTGGCATGCCCCGACACTGCTCATGCC	5060	Db	6081	CAACGCCCTCGGTCTCTTTGACTCAGTGATGACGGCGCATCTCTGAGAGCTCTTTCT	6140
QY	4969	AGATGGCGAGTGTACTGGGTGACCATGGGCACCAACAGTGCACTCAAGAGTGTGACCA	5028	QY	6046	GGGCACTGGCGCGAGGTGATATACAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCT	6105
Db	5061	TGATAATCAGATCATCACCTTACGGTGGGACCAACAGGAGGCTCAAGAGCGTGTCAAC	5120	Db	6141	GGGCAACGGCGCGCAGGTGTTCTATAGTACGGAACCTGTCCTCAAGTTATCGGAGATCGT	6200
QY	5029	ACAAGGACAGTGTGGCCATGATGACATACCAATGGCAATTCGGCTCTTCTGGCAACAA	5088	QY	6106	CTATGACACACCAAGGTGAGTTTCACTATGACGAGACGCGCAGGCATGCTGAGAGCCAT	6165
Db	5121	GCAGAACTGGAGTGGGCTCTGACTTATGATGGGAACACTGGACTCTTAGGCCACCAA	5180	Db	6201	CTACAGACACTGCGCTCACCTTCGGCTATGACGAGACCACTGGCGCTCTGAGATGGT	6260
QY	5089	AAGCAATGAACCGATGGACACATTTATGATGATGACGAGCTTTTGGCGGCTGACAA	5148	QY	6166	CAACCTCAGAAATGAGGCTTCACTGCAACCATCCGCTACCGTCAAGTGTGGGCCCTGAT	6225
Db	5181	GAGCGATGAACCGGATGGACAACTTTTATGATGATGACCAAGCGGCGCTGACCAA	5240	Db	6261	GAATCTCCAAAGCGGGGCTTCTCTGTACCATCAGGTACCGAAGGTTCGGGCGGCTCGT	6320
				QY	6226	TCACCGACAGATCTTCCGCTTCACTGAGGAAGGCATGGTCAACCGCCGCTTTTGATACA-	6284

Db 6321 GGACACAGATTTACAGGTTCTCTGAGGAGGAGCATGATCAACGCCAGGTTGATTTACAC 6380
 Qy 6285 --ACTATGACAAAGCTTCGGGTGACAGCATGCGGTGATGATCAACGAGACCCACT 6342
 Db 6381 CTACACGACAAAGCTTCGGGTGACAGCATCAAGCCCGTTCATCAGTGAGACTCCCT 6440
 Qy 6343 GCCATTGATCTCTATCGCTATGATGATGTCAGGCAAGACAGAGAGTTTGGGAAGTT 6402
 Db 6441 TCCGTTGACCTCTACCGCTACGATGATGATTTCTGCAAGGTGGAACATTCGGCAAGTT 6500
 Qy 6403 TGGTGTCAATTTACTATGACATTAACAGATCATCACAGCTGTGATGACCCACACCAA 6462
 Db 6501 CGGGGTTCATCTACTACGACATCAACAGATCATCACCTGCGCTCATGACACTCAGCAA 6560
 Qy 6463 GCATTTTGATGATATGCGAGGATGAGGAGTGCAGTATGATCTTCGGTCCGTCAT 6522
 Db 6561 GCATTTTGACCCATGGCGCATCAAGGAGTGCAGTATGATGATTTCCGTTCCCTCAT 6620
 Qy 6523 GTACTGATGACCGTCCAGTATGATTAACATGGGCGAGTAGTGAAGAGAGCTGAAGT 6582
 Db 6621 GTACTGATGACGGTGCATATGACAGTATGGCAGGTCATCAAGAGGGAATGAAACT 6680
 Qy 6583 AGACCTTACGCCATACCACTCGCTACTCTATGATGATGATGCTGACGGCCAGTGA 6642
 Db 6681 GGGGCCCTATGCCAAACCAACCAAGTACACCTATGACTACGCGGGGACGGCAGCTCCA 6740
 Qy 6643 GACAGTCTCCATCAATGACAAAGCCACTTGGCGCTACAGTACCACTCAATGGGAACCT 6702
 Db 6741 GAGTGTGGCGTCAATGACCGGCTACCTGGCGTTATAGCTATGACTCAATGGGAACT 6800
 Qy 6703 GCATTTACTGAGCGCTGGGAACAGTGCACGGCTCACACCACTACGTTGATGATCGCGA 6762
 Db 6801 GCACCTGTAACCCAGGAACAGTGTGCGCTCATGCGCTTACGCTATGACCTCGCTGA 6860
 Qy 6763 CCGCATCACTCGGTGGTGCAGTGCATACAGATGAGATGGATGGCTTCTCTAGGCA 6822
 Db 6861 CCGGATCAACAGGCTAGGGAAGTGCAGTACAAATCGATGATGCTATTTATGCA 6920
 Qy 6823 GCGGGCGGTGATCTTTGAGTACAACTCAGCTGCGCTGCTCATCAAGCGCTTACAAACG 6882
 Db 6921 GAGAGGATCTGACATCTTGAATACAACTCAAAGGCTCTTAACGAGAGCGTACAAAC 6980
 Qy 6883 GCGTGCAGCTGAGTGTACGCTACGCTACGCTGCGCTGGGCGCGGCTGTCAGCAA 6942
 Db 6981 GCGCAGCGGTGAGTGTGAGTACCGCTATGATGCGGTGAGCGCGCGGCTTCTTACAA 7040
 Qy 6943 GAGCAGCCACAGCCACCTGAGTTCCTTATGACAGCTGACCAACCCCAACAGGT 7002
 Db 7041 GACCAACTTGGCCACCACTACAGTACTTCTATTCGACCTCCACCCACACAGTAT 7100
 Qy 7003 CACCACTGTACAAACCTCAGCTCTGAGTACCTCCTCTACTACGACTTGAAG 7062
 Db 7101 CACCCATGTTTAAACCACTCCAACTGAGTACCTCACTACTACTGACCTCCAGG 7160
 Qy 7063 ACACCTTTTGCAATGAGCTGAGCTGATGATGATTTTACATAGCTTGTGAACAAT 7122
 Db 7161 CCACCTTTTGCAATGAGAGAGTGTGGGAGAGTACTATGTTGCTCTAGATTAAC 7220
 Qy 7123 CGGACCCCTCTGCTGCTTTTATGGAACAGGTTTGTATGATCAAGCAAACTCTGTAC 7182
 Db 7221 CGGACTCTCTGCTGCTTTTATGATCAATGGCTCATGATCAAGCAAACTCCAATAC 7280
 Qy 7183 AGCTATGGGAGATCTTACATGATGATCAACCCCACTTTCAGATCATATAGGCTTACCA 7242
 Db 7281 AGCTATGGGAGATTTTACTATGATCTCAATCCAGCTTTTCAAGGCTCATCGGCTTCA 7340
 Qy 7243 TGGTGGCTCTATGATCACTACCAAGCTTGTCCACATGGCGCGGAGATTTATGAT 7302
 Db 7341 CGGAGGCTCTACGACCCCTCACCAGCTCGTTTCACTTTACGACGCTGATTTAGCT 7400
 Qy 7303 GCTTGGCGGAGCTGAGCTAGCCACAGCTGCGAGCTGCGAGCTTTAGTAGCAAA 7362

Db 7401 GCTGGCAGGACGGTGGAGCTCCCGGACTACACCATGTGGAGGAA---TGTGGGCAAGGA 7457
 Qy 7363 CGTCATGCTTTTAACTCTATATGTTCAAAACAAACCCCATCAGCACTCCCAAGGA 7422
 Db 7458 GCCAGCCCCCTTCAACCTGTATCTTTCAAGAACAAATCCACTCAGTATAGCTGGA 7517
 Qy 7423 CATCAAGTGTCTTATGACAGATGTTAAACAGCTGCTCTCACCTTTTGGATTCAGCTACA 7482
 Db 7518 TTTAAAGAACTACGTGACAGACGTTGAAGCTGCTCGTGTGTTTGGATTTCAAGCTCAG 7577
 Qy 7483 CAACGTGATCCCTGTTTATCCCAACAGACATGATGCCATGGAACCTTCTTACAGCT 7542
 Db 7578 CAACATCATTTCTGGATTTCCCAAGAGCCAAATGATTTTGTGCTCCCTCCCTATGA--- 7634
 Qy 7543 CATCCACACAGATGAAACCGCAGGAGTGGGACCAACAGCAAGTCTATCTCCGGGTACA 7602
 Db 7635 -----ACTGTACAGAGCCCAAGCAAGTGAGATGGACAGCTCATTACAGGTGTCCA 7685
 Qy 7603 GTGTGAAGTACAGAAAGAGCTCAAGGCTTTGTCTACCTTAGAACGTTTACAGCTCTTA 7662
 Db 7686 GCAGACAAACAGAGAGGATTAACAGGCTTTTGGCTCTTAGAAGGACAGGTCTATCTAA 7745
 Qy 7663 TGGCTCCACATCAACAGCTGCGCAGAGGCTCCAAAGACCAAGAGTTTGCATCCAGCG 7722
 Db 7746 AAAGCTCCATGCGAGGCTCCGAGGAAAGC-----AGGCCACTGTTTCTAGACCAC 7799
 Qy 7723 CTCAGTCTTTGGCAAGGGGTCAAGTTTGGCTTGAAGATGGCGAGTGGCGAGTACCAAGCAT 7782
 Db 7800 GCCCATCATCGCAAGAGCATGTTCCGCATCAAGAAAGGCGGTGACCAAGCGGT 7859
 Qy 7783 CATCAGTGTGGCAATGAGGATGGCGAAGGTTGCTGCCATCTTGAACCATGCCACTA 7842
 Db 7860 GTCTAGCATGCCAGTGAAGACAGCGCAAGTAGCATCCGTGTGTAACCAACGCTACTA 7919
 Qy 7843 CCTAGAAACCTGCACTTCCACATTTGATGGGTGATACCCATTTCTTGTGAACCCAGG 7902
 Db 7920 CTTGGAACAGATGACATACAGCATCGAGGCAAGGACACACATCTTCTGGAAGATCG 7979
 Qy 7903 ACCTTCAGAGGTGACCTGGCCATCTTGGGCTCAGTGGGGGGCGGCAACCTTGGAGAA 7962
 Db 7980 TGCAGCGACGCTGACCTGTTACGCTGGGACCAACCATTTGGGCGCAAGGTGCTGGAGAG 8039
 Qy 7963 TGGGTCACAGCTCTGTGTCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTA 8022
 Db 8040 CGGGGTGAACGTGACCGGTGTCACAGCCACGCTGCTGTAACGAGGACTCGAAGTT 8099
 Qy 8023 CACAGACATCCAGTCCAGTACGGGGCACTGTGCTTGAACACAGCTACCG----- 8073
 Db 8100 CACCAACATTTGAATTTCCAGTACTCCAGCTGCTCAGCATACGCTACGCGCTCACCCC 8159
 Qy 8074 GACAACTTGGATGAGAGAGGACACGGGTCTCGAGCTGGCCCGGAGAGCGCTGCG 8133
 Db 8160 CGACACACTGGATGAAGAGAGGCGCGCTCTGACCAAGCGGACAGAGGCGCTTGG 8219
 Qy 8134 CCAAGCGTGGGCGCGCAGCAGAGACTGCGGGAAGGGGAGAGGCGCTGCGGCGCTG 8193
 Db 8220 TACTGCTGGCCAAAGAGCAGCAGCAAGCCAGGACGGGAGAGGCGAGCGCTGTG 8279
 Qy 8194 GACAGAGGGGAGAGCAGCAGGTTGTCAGACAGGCGGGTGAAGGCTTACAGCGCTT 8253
 Db 8280 GACGAGGGGAGAGCAGCAACTCTCTGAGCAGGCGGCTCAAGGTTATGAGGCTA 8339
 Qy 8254 TTTCTGATCTCTCTGAGCAGTACCAGAACTGTGACAGCGGCCCAACCAATCCACTT 8313
 Db 8340 TTAGTCTCTCCGTGGAACAGTACCAGAGCTGGCAGACAGTAGCAGCAATCCAGTT 8399
 Qy 8314 CATGAGCAGAGCGAGATGGGCGGAGGTGACAGA 8348
 Db 8400 CTTAAGACAGAAATGAGATGGGAAAGAGGTAAACAA 8434

```

: Sequence 66, Application US/09800198
: Publication No. US20030087816A1
: GENERAL INFORMATION:
: APPLICANT: Vernet, Cornie AM
: APPLICANT: Fernandes, Elma
: APPLICANT: Shimkets, Richard A
: APPLICANT: Herrmann, John L
: APPLICANT: Majumder, Kumud
: APPLICANT: Mishra, Vishna
: APPLICANT: Mezes, Peter S
: APPLICANT: Rastelli, Luca
: TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
: FILE REFERENCE: 15966-697
: CURRENT APPLICATION NUMBER: US/09/800,198
: CURRENT FILING DATE: 2001-03-05
: PRIOR APPLICATION NUMBER: 60/186,596
: PRIOR FILING DATE: 2000-03-03
: NUMBER OF SEQ ID NOS: 98
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 66
: LENGTH: 8689
: TYPE: DNA
: ORGANISM: Rattus norvegicus
US-09-800-198-66

```

Query Match	40.7%;	Score 3397.8;	DB 10;	Length 8689;
Best Local Similarity	64.8%;	Pred. No. 0;		
Matches 5469;	Conservative	0;	Mismatches 2712;	Indels 254;
Gaps 19				

35	ATGACGCTGAAGGAGGAGGAGCCTTTACGCGCTCGCTGACCCGCGCGCCGCGACGCCGAGCGC	94
133	ATGGATGTGAAGGATTCGGCGCATCGCTCTTTGACACAGGGGACGGTGTGGCAAGGAGTGT	192
95	CGCTACACGAGCTCGCTCGCGGACAGCGAGGAGGGCAAAG---CCCGCGAGAAATCGTAC	151
193	CGCTACACCGAGCTCTCTCTGGACAGTGGAGACTGCCGTGTGCCACGCGAGAGTCTCTAC	252
152	AGCTCCAGGAGACCTGAAGCCCTACGACAGAGACGCCCGCTAGCCTATGCGACCGCG	211
253	AGTTCCAGTGAAGCCCTGAAGCCTTATGACCATGACAGCAGATGCACTATGGAACCGGA	312
212	GTCAAGGACATTGTCCCGCAGGAGCCGAGGAATTTCTGCCGCACAGTGCACACTTCACCG	271
313	GTCAAGACCTGGTCACCGGAGTCCGATGAGTTTTCTAGACAAGGGGCTAATTTTCACCG	372
272	CTGCGGGAGCTGGGGCTGGAGAGTAACTACGCCCTCACGGGACCCCTGTACCGGACGAGAC	331
373	CTGGCAGAAATGGGAATCTGCGAG---CCCTCCCAACACCGAGTGGTTACTGTTCCGAC	429
332	ATTGGCTGCCCCAAATGCGCTACTCCATGCGGGCTGGCTCTGATGCCGACATGGAGGCT	391
430	ATGGGGATCTCCACAGGGGTACTCCCTGAGCACTGGGTCTGATGCGGACTCGGACACCG	489
392	GACACGGTGTCTGCCCTTGAGCACCCCGTGGCTGTGTGGGGCGGAGCACACGTCACGGG	451
490	GAGGGAGGGAATCTCCAGAAATGCCATCAGACTGTGTGGGACGAGGGATAAATCGAGG	549
452	CGCAGCTCTGCTGCTCCAGCGGGCCAAATTCGAATCTCACATCACCGACACCGAGAT	511
550	CGCAGCTCTGGTGTGTCAGCGGGAGAACTCAGCCCTTACTCTGACTGATTTGCACAAT	609
512	GAAGAACCTGAGACTGATCATCCGGGGGCGCC---TGCAGAAACCAACGCGCGGCTCCGGAC	567
610	GAAGATAAATCGGATGAGCAATGGTTCGACCCCATTCACCTACATCTCTGCTAGCCCTC	669
568	GCGCGCGCGCGCTCTCGCAGCGGCCACACCCCAACAGCACCACCGCGCTCCCAATTA	627
670	CTCCCATCTGCTCAGCTGCTAGTCCCATTAATCTCCACAGTTAGTGGCC---AGATGCC	728
628	CTCCCTGAACCGGGGCAACTTCACGCGAGGAGCAACCCAGCGCCGG-----	674
729	ATTGTTAGACAGCAACACTCCCATCAGATCATGGAACACCAACCCGATGAGGAATCTTC	788

QY	675	-----CCCCACGGACCACTCGTCTCTCGGAGAGCCCTCTCGCGGGCGGCC	722
Db	789	CCCTAAATTATACCTGTCTAGAGCATGCTCAGGCCCCCAGCAAGCTCTCCAGTAGTGGCCC	848
QY	723	AGGAGCTGCCCGCCAGCGAGAACTGGGTCTCAACAGCAACATCCCCCTGGAGACCA	782
Db	849	TCCGAACACACAGCCAGTCAACGTGAGGCCCTCTCTGCCACTCTCATAAACACAC	908
QY	783	GGAACTTAGCAAGCAGCCATTCTTAGGACATTGAGGACAACTCATTTGAGATGGACA	842
Db	909	CCTGTCCACGACCACTCTCTGCACTCCCTCAACAGAAACTCACTGACCAATCGGCG	968
QY	843	TTCTCGGGCGCTCCGGCA-----	861
Db	969	GAGTCAAAATCCACGCCCCAGCTCTCTGCACCCAAATGACCTGGGCCACACGCGCGAGTCCGT	1028
QY	862	-----TGATGGGGCTTACAGTGAAGGGGCACTTCCT	891
Db	1029	TCAGTCCAGGACAGCTGGGTGTGAACGTAACGTGCGCTGGAGACGGGCACTTCCT	1088
QY	892	CTTTCAA---GCCTGGAGGCACTCCCGGTCTTCTGCACCAATACACAGGTACCCTACT	948
Db	1089	CTTCAAGACGTCTCGGAAGCACACCCCTGTTTCAAGAGCTCTTCTTCCAGGATACCCCTT	1148
QY	949	GAGTCCAGCACAGTGTACTCTCTCGCGCCCGACCCCTGCCCGCAGACACCTTTCGCCCG	1008
Db	1149	GACCTCAGGACCGTTTATACACCAACACCCCGCTGCTGCCACGGAATACATCTTCTAG	1208
QY	1009	GCGGCCCTTAACTCAAGAAAGCCCTCCAAAGTACTGTAACTGGAAGTGGCGACCTGAG	1068
Db	1209	GAAGGCCCTTCAAGCTGAAGAAACCTCCAAATACTGCAGTTGGAAATGGCGCCCTGTC	1268
QY	1069	CGCATCGTCTCATCTCAGCCACTCTGTGTCTCTGCGCATCTTTGTGGCCATGCACCT	1128
Db	1269	TGCCATTGCCGTGCCCTCTCTTGCCCATTTTGTGGCCATTTTCATAGCAATGCATCT	1328
QY	1129	GTTTGGCCTAAACTGGCACTCGACCGATGGAGGGCAGATGTATAGATACGGAGGA	1188
Db	1329	GCTCGGACTCAATTGGCACTCCAGCGCGAGATGCACACACCCTTTAAACA-----ATGG	1382
QY	1189	CACAGCCAGCTGGCTGTGCCAACCGGACGTCTCCCTATACCCCTCAGGGGGCACTGG	1248
Db	1383	CGTAAGACCGGCTTACAGAAACGATGATGGCAACAGTGCCTCTCTGGAGGCAAGT	1442
QY	1249	CTTAGACACCCCTGACAGGAAAGCAAGAAACACAGAAAGAAAGCCAGTAGTTTCTT	1308
Db	1443	GCCCTGCTGTGAAAA-----	1459
QY	1309	TCAGAGGACAGTTTTCATAGATTCTGGAGAAATGTATGGGAAGCGGAGCCTCCAGAA	1368
Db	1460	-----ACAGCAGCATACACGCGCGAGGAGCAAGTGGTCTGCGGATGCACAGGA	1511
QY	1369	GATTCTCTCTGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCACTGAA	1428
Db	1512	AGTCCACACAGGGTGTTTTGGAGTCCAGATTTCATCATGTCAGCCTCAGTTCTTAA	1571
QY	1429	ATTCAATGTGTCTCTGGAAAGCAGCCCTGTGTGGCATTTATGGCAGAAAGCCCTCC	1488
Db	1572	GTTCAAATCTCCTCGGGAAGGATGCCCTCTTCGGCGTCTACATAGAAGAGGACTGCC	1631
QY	1489	TCCTTCATATACAGTTTGCATTTGTGGAGCTGTGGATGGCAGGAGGCTCTTAACCCA	1548
Db	1632	ACCATCTCATGCAAGTATGACTTTATGAAACCCCTGGACGAA-----A	1676
QY	1549	GGAGGCGGAGCCTTAGAGGGGACCCCGGCCAGTCTCGGGAACTGTGCCCCCTCCAG	1608
Db	1677	GGAGAACTGGAGTGTGTCTGAGTCACCCAGGGAACCGCCGAGCATCCAGACCCCTGGTGA	1736
QY	1609	CCATGACAGGCTTCATCCAGTATTGGATTGAGAACTCGGCATCTGGCTTTTACAA	1668
Db	1737	GAACGAGGCTGTGTCTGTGAGTACTTGGATGTGGGCTGTGGCAGCTCGCTCTACAA	1796
QY	1669	TGACGGAAGAGGAGTCAGAACTGTTTCTTTCTTCACCACTGCATTAGTTCGGTGGATAA	1728

1797 TGACGGCAAGGACAGAGATGCTCTCTTCAATACGGTGTCTTAGATTCAAGTGCAGGA 1856
 1729 CTGCCCCAGCAACTGTCTATGGCAATGGTGAATGCTCTCTGAGACCTGCCATCTCTCT 1788
 1857 CTGTCCACGAAATGCGACAGGAAACGGCGAATCGGTCTGGAATGCTCTCTCTCTCT 1916
 1789 GGGTTTCTGGGGCCCCGACCTGTGGCAGAGCTCTGCCCCGCTCTCTGTAGCGGAAATGG 1848
 1917 AGGATTCCTAGGTGCAGACTGCGCTAAAGTGCCTGCCCTGTCTGTGAGTGGGAATGG 1976
 1849 CCAATACATGAAGGCGAGATGCTTTGTGCAAGTGGCTGGAAGGCGCTGAGTGGATGT 1908
 1977 ACAGTATTCCAAAGGACATGCGCAATGCTCTACAGTGGCTGGAAGGAGCAGAAATGGATGT 2036
 1909 GCCCACCACCAAGTGTATCGATGTGGCTGCACCAACCATGGCACCTGCATCACGGGCAC 1968
 2037 GCCCATGAACAGTGCATCGATCTCTCTGTGGGGCCACGGCTCTGCAATGATGGGA 2096
 1969 CTGCATCTGCAACCTGGCTTACAGGGCGAGAGCTGTGAGGAAGTGGACTGCAATGGAACC 2028
 2097 CTGCGTGTGTCAGCTGGCTTACAGGGCGAGACCTGCAAGAGAGTGGATTTGCTTGGATCC 2156
 2029 CACATGTTACGGGGGGTGTCTGCTGAGAGGCGAATGCCATGCTTTGTGGGATGGGG 2088
 2157 AACCTGCTCCAGCCATGGTGTCTGTGGAACGGAGAGTGTCTAAGCAGCCCCGGCTGGGG 2216
 2089 AGGCACCAACTGGAGACCCCCAGGGCCACATGCTTTAGACCAAGTGTAGGGCCACGGAA 2148
 2217 CGGGCTCAACTGCGAGTGGCGAGGGTCCAGTCCCCAGACCAAGTGTAGTGGGCATGGCAC 2276
 2149 CTTCCTCCGGACACCGGGCTTTGAGCTGTGACCTGACCCCAAGCTGACGACGACTGTT 2208
 2277 TTACCTCCCTGACTGTGGCTCTGCAACTGTGATCCGAAATGGATGGGTCCCGACTGTC 2336
 2209 TATCGAGATCTGTCTGCCACTGTGGTGGCCATGGCTGTGGTAGGGGGACCTGGCG 2268
 2337 TGTGAAGTGTCTAGTACATGTGGCACTACAGCGCTGTGATCGGGGGAGCTGGCG 2396
 2269 CTGCGAGGATGGCTGGATGGGGGAGCTGTGCAACAGCGGGCTGCAACCGCGCTGTGC 2328
 2397 CTGTGAAGAGGGTGGACAGCGCGGCTTGTGACAGCGGCTGTGACACCGCGCTGTGCAT 2456
 2329 CGAGCATGGGACCTGCGCGACGCGCAAGTGCAGTGCAGCCCTGGCTGGAATGGCGAACA 2388
 2457 TGAGCACGGACCTGTAAAGATGGCAAAATGTAATGCCAGAGGGCTGGAATGGTGAACA 2516
 2389 CTGCACCATCGCTCACTATCTGGATAGGGTGTAAAGAGGGTGGCCCTGGGTGTGCAA 2448
 2517 CTGCACCATTTGAT-----GGCTGCCCTGATTTGTGCAA 2549
 2449 TGCCACGGCAGATGTACCTTAGACCTGAATGTTGGCACTGGCTGCTGCCAGCTGGGCTG 2508
 2550 CGGTAAACGGAGATGCACATGGGTGAGAACAGCTGGCAGTGTGTCTGCCACAGCGGCTG 2609
 2509 GAGAGAGCTGGTGTGACACTTCCATGGAGATGCTCTGGGTGACAGCAAGCAATGA 2568
 2610 GAGAGGGCCCGATGCAACCTTCCATGGAACTCTCTGGCTGATAACAAGGATAATGA 2669
 2569 TGAGATGGCTGGTGAATGACCTGTGACCTGTGCTGCTGCCAGCCCTGTGCCATAT 2628
 2670 GGGAGATGGCTGGTGAATGACCTGTGACCTGTGCTGCTGCCAGTGTGTCTGCCAGCGGCTG 2729
 2629 CAACCCCTGTGCTGGCTGCCCTTAACCTCTGGACATCATCCAGGAGACACAGTCCC 2688
 2730 CAGCTGCTGTCTGGGGGTCTCGGACCCCTTTGGACATCATTCACAGGCGCAGACAGA 2789
 2689 TGTGTACAGAGAACCTCACTCTCTTATGACCGGATCAAGTTCCTGTGGGCGAGGA 2748
 2790 CTGGCTTCGGGTGA-----AGTCTCTTATGATCGTATCAAGCTCTTGGCAGGCAAGGA 2843
 2749 CAGCAGCACATAATCCCGGGGAGAACCCCTTTGATGGAGGCGATGCTGTGTTATTCCG 2808

2844 CAGCACCCACATCATCTCTGGAGCAACCCCTTTCAATAGCAGCCTGGTGTCTCTGTATCCG 2903
 2809 TGGCCAAAGTATGATCATCAGATGGAACCCCTCTGGTGTGTGAACATCAGTTTGTCAA 2868
 2904 AGGCCAAGTAGTAACACAGGATGGACCCCTCTGGTGTGTGAATGTGTCTTTGTCAA 2963
 2869 TAAACCTCTCTTTGGATATACATCAGCAGGCAAGATGGCAGCTTTGACTTTGGTGTGACAA 2928
 2964 GTACCCAAAATATGGCTACACCATCACTGCCAGGACGGACGCTTTGACCTGATTTGCCAA 3023
 2929 TGGCGGCATCTCCATCATCTGGGTTTCGAGCGGACCTTTTCATCACAGAGGACAC 2988
 3024 TGGGGCTCTGCGCTGTGATCTCTTCACTTTGAGCAGGCCCTTTTCATGACCGGAGGCGAC 3083
 2989 CCTGTGGCTGCATGGGATCGCTTTCTTTGTCTATGGAACCATCATCATGAGACATGAGGA 3048
 3084 AGTATGCCCGCGTGGAAACAGCTTCTATGCCATGGACACCCCTGGTAATGAAGACGGAGGA 3143
 3049 GAATGAGATTCAGAGCTGTGACCTGAGCAATTTTGGCGGCCCCAACCCAGTCTCTCTCC 3108
 3144 GAATCCATCCCCAGCTGTGACCTCAGTGGCTTTGTTCGGCCTGATCCGATCATCATCTC 3203
 3109 ATCCCCACTGACGTCCTTCCGCCAGCTCCTGTGCAGAGAAAGGCCCATTTGTGCGGAAAT 3168
 3204 CTCTCTCTGTCCACTTCTTCAGCGCTTCCCTGGCGCAACCCCATTTGTGCTCTGAGAC 3263
 3169 TCAGGCTTTGAGGAGAAATCTCTATCTCTGTGCTGCAAGATGAGGTGAGCTTACCTGAG 3228
 3264 CCAGGTTCTTCATGAGGAGATCGAGCTCCCTGGCACCAACCTGGAAGCTCCGTTACCTCAG 3323
 3229 CAGCGGACCCCTGGCTGACAAATCTGCTGAGGATCAGCTCACCCACCCGACCATCCC 3288
 3324 CTCAGAACAGCAGGATCAAGTCACTGCTGAGATCACCATCACCCAGTCTCAGGTGCC 3383
 3289 CTTCAACCTCATGAAGTGCACCTCATGTTAGCGGTGAGGGCGGCTCTTTCAGGAAGTG 3348
 3384 CTTGAACCTCATCCGGTTCACCTGATGTTGCGTGGAGGGGCTCTCTTTCAGGAAGTC 3443
 3349 GTTCGCTGCAGCCCCAGCTGTCTTATTTTCATTTGGGACAGACACAGCTTCTACAA 3408
 3444 GTTCCAGCTTCTCCCACTGGCTTACATTCATCTGGGACAGACAGCTTATGG 3503
 3409 CCAGAGGTGTTTGGGCTTTTCAGAGGCTTTGTTTCGCTGGGTTATGAATATGAATCTCTG 3468
 3504 CCAGAGGTTTATGGCTATCGGATGCTGTTGTGCTGTTGGATTTTGAATATGAGACCTG 3563
 3469 CCCAGATCTAATCTGTGGGAAAAAGAACAAAGTCTGCTGAGGGCTTATGAATTTGACGC 3528
 3564 CCCAGTCTCATCTGTGGGAAAAAGACAGCCCTACTTCAAGGATTCAGGCTGGACCC 3623
 3529 GTCCAAAGCTTGGAGGATGGAGCCTTAGACAAACATCATGCCCTCAACATTTCAAAGTGGTAT 3588
 3624 TTCCAACTTGTGGTGGTCCCTGGATAGCACCAACCTCMAATGTGAAGAGCGGAAT 3683
 3589 CTTGCACAAAGGGAATGGGAGAACAGTTTGTGTCTGACAGCCTCTCTCATTTGGGAG 3648
 3684 ACTACTCAAAGGCACAGGGAGAACCAAGTTCTCTGACCGAGCCGCCCATCATCACAG 3743
 3649 CATCATGGCAATGGGCGCGGAGAACATCTCTGCCCGCAGCTGCAACCGCTTGTCTGA 3708
 3744 CATCATGGTAAACGGTGGCGGAGACATCTCTCTGCCAGCTGCAATGGCTTGTCTGA 3803
 3709 CGGCAACAAAGCTCTGCGCCCGAGTGGCCCTCACTGTGGTCTGACGGGAGCTCTTATGT 3768
 3804 AGGCAACAAACTGTGGCCCGGCTGGCTGTGGGATCGATGGGAGCTCTTGT 3863
 3769 GGGTGAATTTCACTACATTAAGAGATCTTCCCTCTGGAAATGTCCACCAATCTCTAGA 3828
 3864 CGGTGACTTCAATTAATATCGGGGCACTTCTCCCTTCTCGAAACGTGACCAAGTATCTTGA 3923
 3829 GCTGAGGAATAAGATTTTCAGACATAGTACAGTCCAGCACACAAATFACCTCTGGCCAC 3988
 3924 GTTACGAAATAAGAGTTTAAACATAGCAACAGCCAGGACACAAAGTACTACTTGGCTGT 3983

Qy 3889 AGACCCCATGAGTGGGCGCTCTTCTTTCTGACGACCAAGCGCGGCTCTTTAAAT 3948
 Db 3984 GGACCTGTGACTGGCTCGCTATGCTCTGACACCAACAGTCGCGGATCTACCGAGT 4043
 Qy 3949 CAAGTCACCTGTTGGTGTGAGGACCTTGTCAAGAACTCTGAGAGTGGTTCGGGGACAG 4008
 Db 4044 CAAGTCTCTAAGCGGAGCCAAAGACCTGGCTGGGAATTGGAAATTGTGGCGGGACTGG 4103
 Qy 4009 TGACCACTGCTCCCTTTGATGACACTCGCTCGGGGATGGTGGGAAGGCCACAGAAGC 4068
 Db 4104 CGAACATGCTACCTTTGATGAAGCCGCTGTGGGATGGCGGAAGGCTGTGATGC 4163
 Qy 4069 CACACTCACCAATCCGAGGGGTATTACAGTGGACAAAGTTTGGGCTGATCTATCTTCTGGA 4128
 Db 4164 CACCTGATGAGCCCTAGAGGTATTGAGGTAGCAAGAACTGGCTGGGAATTGTGGCGGGACTGG 4223
 Qy 4129 TGGCACCATGATCAGACGCACTCGATCAGAAATGGATCATCTCCACCTCGTGGCTCTAA 4188
 Db 4224 TGGCACCATGATCGGAGCTCGATGAGCCGCTGTGGGATGGCGGAAGGCTGTGATGC 4283
 Qy 4189 TGATCTCAATACAGCCGCGCACTCAGCTGTGATCTGTCAAGGATATTTCAGGTAAG 4248
 Db 4284 TGACCTCACAGCTGTCCGACCACTGAGCTGTGACTGTAGCATGAGCGTGGCCAGGTCCG 4343
 Qy 4249 ACTGGAGTGGCCACAGACTTAGCCATCAACCAATGGCAACTCACTTATGTCTCGA 4308
 Db 4344 TCTAGAAATGCCGACAGACCTTGGGTCAACCCCATGGCAATTCCCTGTAGCTCTGGA 4403
 Qy 4309 CAACAATGTGGTCTCTGAAATCTCTGAAACCAACCAAGTGGCAATTGTGCCGGAGGCC 4368
 Db 4404 GAACAACGTCTCTCGGATCACCGAGATCACAGGTCAGCATCATCGCGGACGGCC 4463
 Qy 4369 CATGCACTGCGAGTCCCTGGATTTGACCACTTCTGCTTAAGCAAGTGGCCATCCACGC 4428
 Db 4464 CATGCACTGCGAGTTCGGGCTCGCACTACT--CGCTCAGCAAGCTCGCCATCCACTC 4520
 Qy 4429 AACCTGGAGTCAGCAACCTTGGCTGTTTCAACAATGGGCTCTGTATATCTCTGA 4488
 Db 4521 TGCTCTGGATCAGCAAGCCCATCGCCATTTCTCACCGGGGTGCTCTACATCACGA 4580
 Qy 4489 GACTGATGAGAAAAGATCAACCGCATCAGGCAAGTCAACACTAGTGGAGATCTCAT 4548
 Db 4581 GACGGACGAGAAGATCAACCGCTTACCGCAAGTCAACCAACGAGAGATCTGCCT 4640
 Qy 4549 CGTGTCTGGGCCCCAGTGGCTGTGACTGTAAATGATGCCAATGTGATTTTTC 4608
 Db 4641 CTTAGCCGGGAGCTCTGACTGTGACTGCAAAATGAGCTCACTGATCTGCTATTTC 4700
 Qy 4609 TGGAGACGATGGTTATGCCAAGGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGTG 4668
 Db 4701 GGGAGATGACGATACGCCACGATGCCATCTTTGAACTCCCGCTCTCTAGCTGTGC 4760
 Qy 4669 TGCTGATGGGAGCTCTAGTGGCGACCTTGGAACTCCGAAATCGGTTTATCCGGAA 4728
 Db 4761 TCCGGATGGCACCATCTACATCGACACCTCGGGAATATCCGATCAGGCGGTACGAA 4820
 Qy 4729 GAACAAGCTTTCTCAACACCCAGAACATGTATGAGCTGTCTTCAACCAATTCACCAAG 4788
 Db 4821 AAACAACCTTGTCTTAAACCGTTCAACAGATGAGGCTGCTCTCCGGAGAACAGGA 4880
 Qy 4789 GCTCTATCTGTTGATACCGGCAAGCACTGTATACACCAAGCTGCCACAGAGA 4848
 Db 4881 ACTGTACGTGTTCAACCGCGATGGTATCCATCAGTACACCGTGTGAGCTGGTACCGGGA 4940
 Qy 4849 CTACCTGTACACTTCACTTACACTGGGACGGGACATCACAATCATCACAGAACAA 4908
 Db 4941 GTACTTATCAATTTCACTTACAGCGTGTGAATGATCACCGAGTTGATTGACAAACAA 5000
 Qy 4909 TGGCAACATGGTAAATTCGCGGAGACTCTACTGGGAATCCCTCTGGGTGGTGGTCCC 4968
 Db 5001 CGGGAATTCCTTAAAGATCCGCGGACAGCAGTGGCATGCCCGACACCTGCTCATGCC 5060

Qy 4969 AGATGGCCAGGTGTACTGGGTGACCATGGGCACCAACAGTGCATCTAAGAGTGTGACAC 5028
 Db 5061 TGATAATCAGATCATCACCTTACGGTGGGCACCAACGAGGCTTCAAAGCGGTGTCAAC 5120
 Qy 5029 ACAAGGACACAGTGGCCATGATGACATCAATGGCAATTCGCGCTTCTGGCAACCAA 5088
 Db 5121 GCAGAACCTGGAGCTGGCCCTCATGCTTATGATGGGAACACTGACTCTCTAGCCACCA 5180
 Qy 5089 AAGCAATGAAACCGGATGGACCAATTTATGAGTACGACAGCTTTGGCCCTGTGACAAA 5148
 Db 5181 GAGCGATGAAACCGGATGGACCAATTTTATGACTATGACCAACGAGGCGCTGTGACCAA 5240
 Qy 5149 TGTGACCTTCCCTACTGCGCAGGTGAGCAGTTTCGAAAGTATACAGACAGTTCAGTCA 5208
 Db 5241 TGTGACTGCGCCCAACGGGGTGGTACAGCCTGACCGGAAATGGAGAAATCCATCAC 5300
 Qy 5209 TGTCCAGGTAGAGACTCCAGCAAGGAT--GATGTACCAATAACCAACCAACCTGTCTGC 5265
 Db 5301 CGTTGACATTGGAACCTCAACCGGTGATAACGATGTCACTGTGATTACCAACCTCTCTTC 5360
 Qy 5266 CTCAGGCGCTTTTACACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGC 5325
 Db 5361 AGTGAGGCGCTCTACACCGTGTACAAGATCAAGTGCAGAACAGCTACCAAGCTCTGAG 5420
 Qy 5326 CGATGGCTCTTGGCGCTGCTGGCCAAACGGCATGGAGTGGCGCTGCAGACTGAGGCC 5385
 Db 5421 CAAACGGACCTTGGCGCTGATGACGCAACGGCATGGCGCTGAGCTTCCACAGGAGCC 5480
 Qy 5386 CCATTTGTGCTGTCACCTGCAACCCACCGTGGCAAGAGGAATGTCAAGCTGCCCAT 5445
 Db 5481 CCACGCTCTCGAGGACCTCTCACCCACCATCGGGGCTGTAACATCTCCCTGCCCAT 5540
 Qy 5446 CGACAAACCGCTCAACTGTTGGATGGCGGACGCAAGAGCAGGCTCGGGCCAGGT 5505
 Db 5541 GGAGAACCGCTGAATCCATCGAGTGGCGCTGAGGAAGAACAGATTAAGGCAAGT 5600
 Qy 5506 CACTCTCTTTGGCGCGCTCGCGGTGCACAAACCGAAATCTCTATCTCTGGAATTTGA 5565
 Db 5601 CACCATCTTTGGAGAGACTTCGGTCCACGGAAGAACTCTCTGTCTCATGATTATGA 5660
 Qy 5566 TGCGGTACACGACAGAGAAATCTATGATGACCAACGCAAGTTTCAACCTTCGATTC 5625
 Db 5661 CCGAAATATCGCATGTAGAAAGATCTATGACGACCAACCGGAAGTTTCACTCTGAGGATCAT 5720
 Qy 5626 GTACACACAGGCGGGCGGCCAGCTCTGTGTACCCAGCAGCAGGCTGAATGGTGTCAA 5685
 Db 5721 TTATGACAGGTGGCGCGCTTCTGTGCTCCCCAGCAGTGGACTGGCGCGCTCAA 5780
 Qy 5686 CGTGACATCTCCCTGGGGTTACATTTGCTGGCATCCAGAGGGCATCATGTCTGAAAG 5745
 Db 5781 TGCTCTCTACTTCTTCAACGGCGCTGGCGGCTCCAGCGCGGGCCATGAGCGAGAG 5840
 Qy 5746 ATGGAATACGACAGGCGGGCCGCATCACATCCAGGATCTTGGTGTGGAGACATG 5805
 Db 5841 GACAGACATTGACAAAGCGCGGATTTGTGTCGGAATTTGCGCGGACGGGAAAGCTGT 5900
 Qy 5806 GAGCTTACACATCTTATGAGAAAGTCCATGGTGTCTACTACACAGCAGAGGAGTATAT 5865
 Db 5901 GAGCTTATCTTACCTTGAAGTCCATGCTCTCTGCTGCGAGCAGCGCTAGTACAT 5960
 Qy 5866 CTTTGGATTGCAAGAAATGACCGCTCTCTTGTGAGATGCCCAACGTTGGCGCGCA 5925
 Db 5961 ATTTGAATGATGATCTCTTGAACCGCTCCAGCAGTCACTGCGGAGTGTGCGCCGCA 6020
 Qy 5926 GACACTAGAGACCATCCCTCAGTGGGTACTTACAGAAACATCTATCAGCCCTGTAGGG 5985
 Db 6021 CAGCATGTCACGACACTCCATTTGGTATCATCGGAACATTTACACCCACCGGAAG 6080
 Qy 5986 CAATGCCCTGATGATPACAGGACTTCACTGAGGATGGGCACTCTCTTACACCTTCTAGCT 6045
 Db 6081 CAAACGCTCTGGTCTATCTTTGACTACAGTGAAGCGCGCATCTCTGAAGAGCTCTTTCT 6140
 Qy 6046 GGGCACTGGCGCGAGGTGATATCAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCT 6105

Db 6141 GGGCAGCGGCGAGGTGTTCTATAGTAGCGAAACTGTCAAAGTTATCGAGATCGT 6200
Qy 6106 CTATGACACCAACCAAGGTGAGTTTCACTATGACGAGCGGAGGATGCTGAAGACCAT 6165
Db 6201 CTAGCAGCAGCTGCGGTGACCTTGGCTATGACGAGACACTGGCGTCTGAAGATGT 6260
Qy 6166 CAACCTACAGAAAGAGGGCTTCACTGACACCTCCGCTACCGTCAGATGGGCGCCCTGAT 6225
Db 6261 GAATCTCAAAGCGGGGCTTCTCCTGTACCATCAGTACCGAAAGGTGGGCGCCCTGT 6320
Qy 6226 TGACCGACAGATCTCCGCTTCACTGAGGAAGCATGGTCAACGCCGCTTTTGACTACA- 6284
Db 6321 GGACAAGCAGATTACAGGTTCTCTGAGGAAGCATGATCAACGCCAGGTTGATTTACAC 6380
Qy 6285 --ACTATGACAAAGCTTCCGGGTGACCAAGCATGCGAGCTGTGATCAACGAGACCCACT 6342
Db 6381 CTACCCAGACAAGCTTCCGCTCGCCAGCATCAAGCCGCTCATCAGTGAGACTCCCT 6440
Qy 6343 GCCATTGATCTTATCGCTATGATGATGTGTCAGGCACACAGACAGCTTTGGGAAGTT 6402
Db 6441 TCCGTTGACCTCTACCGCTACGATGAGATTTCTGGCAAGGTGGAACACTTCCGCAAGTT 6500
Qy 6403 TGGTGTCAATTACTATGACATTAAACAGATCATCACACAGCTGTGATGACCCACACCAA 6462
Db 6501 CGGGGTCACTACTACGACATCAACAGATCATCACCTGCGTCAAGACACTCAGCAA 6560
Qy 6463 GCATTTGATGATATGCGAGGATGAAGAGTGCAGTATGAGATCTTCGGCTCGCTCAT 6522
Db 6561 GCATTTGACACCATGGGCGCATCAAGGAAGTGCAGTATGAGATTTCCGGTCCCTCAT 6620
Qy 6523 GTACTGATGACCGTCCAGTATGATTAACATGGGCGAGTAGTGAAGAGGAGCTGAAGGT 6582
Db 6621 GTACTGATGCGGTCAATATGACAGTATGGCAGGGTCAACAGAGGGAAGTGAAGT 6580
Qy 6583 AGGACCTTAACCAATACCACTCGCTACTCTATGATGATGATGCTGACGACCTCAATGGGAAGT 6702
Db 6681 GGGGCGCTATGCCAAACACCAAGATACACTATGACTACGAGCGGAGCGGCGAGCTCCA 6740
Qy 6643 GACAGTCTCAATCAATGACAAAGCACTCTGGGCTACAGTACGACCTCAATGGGAAGT 6702
Db 6741 GAGTGGCGGTCAATGACCGGCTTACCTGGGCTTATGCTATGACTCAATGGGAAGT 6800
Qy 6703 GCATTAAGCGGCTGGGAACAGTGCAGGCTCACACCACTACGCTATGACATCGGCA 6762
Db 6801 GCACCTGCTAAACCCAGGAACAGTGTGCTCGCTCATGCGCTTACGCTATGACTCGCTGA 6860
Qy 6763 CCGCATCACTCGGCTGGGTGACGTCGAATACAGATGATGAGGATGGCTTCTGAGGCA 6822
Db 6861 CCGGATAACAGGCTAGGAGCGTGCAGTACAAATCGATGATGATGGCTATTTATGCA 6920
Qy 6823 CCGGGCGGTGATATCTTTGAGTACAACTCAGCTGGGCTGCTCATCAAGGCGCTCAACCG 6882
Db 6921 GAGAGGATCTGACATCTTTGAATACAACTCCAGGCGCTTCTAAOCGAGAGCGTACACAA 6980
Qy 6883 GGCTGGAGCTGAGTGTAGGTACCGCTACGATGCGCTGGGCGCGGCGGTGTCCAGGAA 6942
Db 6981 GGGCAGCGGTGAGTGTGAGTACCGCTATGATGGCGTGAAGGATGATGGCTATTTATGCA 7040
Qy 6943 GAGCAGCCACAGCACCACTCGAGTCTTCTATGACAGCCTGACCAACCCCAACCAAGT 7002
Db 7041 GACCAACTGGGCGCACCACTACAGTACTTCTATTCGACCTCCACCAACCCCAACAGTAT 7100
Qy 7003 CACCCACTGTACAAACACTCCAGCTCTGAGATCACTCCCTCTACTACGATTTGAAG 7062
Db 7101 CACCCATGTTTACAAACACTCCAACTCTGAGATCACTCACTATGACTCCAGG 7160
Qy 7063 ACACCTTTTGCAATGAGCTGAGCTGATGATGATTTTATCATAGCTTTGTGACACAT 7122
Db 7161 CCACCTTTTGCAATGAGAGAGTGTGAGGGAAGTACTATGTTGCTCTCAGATTAAC 7220
Qy 7123 CGGACCCCTCTTCTGCTCTTTAGTGAACAGGTTTGTATGATCAAGCAAAATCTGTATAC 7182

Db 7221 CGGACTCCTCTGGCTGTTTTTAGTATCAATGGCTCATGATCAAGCAACTCCATACAC 7280
Qy 7183 AGCCTATGGGAGATCTACATGGATCAACCCCAACTTTTTCAGATCATATAGCTACCA 7242
Db 7281 AGCCTATGGGAGATTTTACTATGACTCCAATCCAGACTTTTTCAGATGCTCATCGCTTCCA 7340
Qy 7243 TGGTGGCTCTATGATCACTCAACCAAGCTTGGCCATCGGCGGAGATATGATGT 7302
Db 7341 CGGAGGCTCTTACAGCCCTTCAACCTGTACATGTTCAAGAACAATCCAATCAGTATGACGT 7400
Qy 7303 GCTGGCGGAGCGCTGAGTACAGCCAGACCTGTGAAGCACCTTATGAGCAGCA 7362
Db 7401 GCTGGCAGGAGGTGGAGCTCCCGGACTACCATGTGGAGAA---TGTGGCAGGA 7457
Qy 7363 CGTATGCTTTTAAATCTCTATATGTTCAAAACAACCCATCAGCAACTCCAGGA 7422
Db 7458 GCCAGCCCTTCAACCTGTACATGTTCAAGAACAATCCAATCAGTATGAGCTGA 7517
Qy 7423 CATCAAGTGTCTATGACAGATGTTAACAGCTGCTCTCACCTTTTGGATTCAGCTACA 7482
Db 7518 TTTAAGAACTACGTGACAGAGTGAAGAGTGGCTGCTGATGTTTGGATTTGAGCTCAG 7577
Qy 7483 CAACGTATCCCTGTTATCCAAACAGACATGGATGCCATGGAAACCTTCCTACGAGCT 7542
Db 7578 CAACATCATCTCTGGATTCCCAAGAGCAAAATGATTTTGTGCTCCCGCTATGA-- 7634
Qy 7543 CATCCACACAGATGAAGAGGAGTGGGCAACAGCAAGTCTATCCTCGGGTACA 7602
Db 7635 -----ACTGTCAGAGAGCCCAAGCAAGTGAAGTGGACGCTCATTAACAGTGTCCA 7685
Qy 7603 GTGTGAAGTACAGAAAGAGCTCAAGGCTTTGTACCTTAGAAGCTTTTACCAGCTCTA 7662
Db 7686 GCAGACAACAGAGAGGAGGATCAACAGGCTTTTGTGCTTAGAAGGACAGTCTATCTAA 7745
Qy 7663 TGGTCCACATACAGCTGCCAGAGGCTCCAAAGACCAAGAGTTTTCATCCAGCG 7722
Db 7746 AAAGCTCATGACAGGATCCAGAGAAAGC-----AGGCCATGTTTCTACGACCAC 7799
Qy 7723 CTCAGTCTTTGGCAAGGGGTCAAGTTTGGCTTTGAAGATGGCGGAGTACCACAGACAT 7782
Db 7800 GCCCATCATCGGCAAGGCACTATGTTCCGCAATCAAGAGAGGCGGTGACACAGCGT 7859
Qy 7783 CATCAGTGTGGCCTAATGAGAGTGGCGAAGGTTGCTGCCATCTTGAACCATGCCACTA 7842
Db 7860 GTCTAGCATCGCCAGTGAAGGACGCGCAAGTAGCATCCGTTGTGAACAACGCTACTA 7919
Qy 7843 CCTAGAAACCTGCACCTTCAACATGATGGGTGGATACCCATTTCTGTGAACCCAGG 7902
Db 7920 CTTGGACAGATGCACTACAGCATCGAGGCAAGGACACACATCTTCTGTGAGATCGG 7979
Qy 7903 ACCTTCAGAAAGGTGACTGGCCATCTTGGGCTCAGTGGGGGCGGCGAACCCTGGAGAA 7962
Db 7980 TGCAGCGGACGCTGACCTGGTTACGCTGGGACCCACCATTTGGCGCAAGGTGTGGAG 8039
Qy 7963 TGGGCTCAAGCTGCTGCTCCAGATCAACAGTACTTAATGGCAGGAGTACAGCTA 8022
Db 8040 CGGGGTGAACGTGACCTGCTCACAGCCACGCTGCTGGTGAACGCGAGGACTCGAAGTT 8099
Qy 8023 CACAGACATCCAGTCCAGTACGGGCACTGTGTTGAACACACGCTTACCGG----- 8073
Db 8100 CACCAACATTTGAATTCAGTACTCCAGCTGCTGCTCAGCATACGCTACGCGCTCACCCC 8159
Qy 8074 GACAAAGTGTGATGAGAGAGGCAAGGCTCTGAGCTGCGCGGCGAGAGAGCGCTGCG 8133
Db 8160 CGACACACTGATGAAGAGAGGCGCGCTCTGGACCAAGCGCGACAGAGGCGCTTGGG 8219
Qy 8134 CCAAGCGTGGGCGCGCAGCAGACAGACTTGGCGGAGGAGGAGGCGCTGGGGCGCTG 8193
Db 8220 TACTGCTGGCCAAAGAGCAGCAGAAAGCCAGGACGGGAGAGAGGCGCTGTGTG 8279
Qy 8194 GACAGAGGAGAGAGCAGCAGGTGTGAGCAGAGGCGGGTGTCAAGGCTTACAGCGCTT 8253
Db 8280 GACGAGGCGGAGAGCAGCAACTCTCTAGCAGCGGACGGGTCAAGGTTTATGAGGCTA 8339

Qy	3254	TTTCGGTATCTCTGTGAGCAGTACCCAGAACTGTTCAGACAGGCCCAACACATCCACTT	8313
Db	8340	TTACGTGCTTCCCGTGGAACTGACCCAGAGCTGGCAGACAGTAGCAGCAACATCCAGTT	8399
Qy	8314	CATGACACAGAGCGAGATGGCGCCGGAGTGACAGA	8348
Db	8400	CTTAAAGACAAATGAGATGGGAAAGAGGTAAACAA	8434

RESULT 11

```

US-09-808-602-7
; Sequence 7, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herрман, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: Macdougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuciec Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 7
; LENGTH: 9826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (280)..(8478)
; US-09-808-602-7

```

Query Match	40.5%;	Score 3385;	DB 9;	Length 9826;
Best Local Similarity	64.7%;	Pred. No. 0;		
Matches 5393;	Conservative	0;	Mismatches 2790;	Indels 154; Gaps 18;
QY	35	ATCGAGCTGAAGAGAGAGAAACCTTACCGCTCGCTGACCGCGCGCCGACGCGGACGCCG 94		
Db	280	ATGGATGTAAGAGACCGCGACACCGCTCTTTGACCAAGAGACCGCTGTGGCAAGAGTGT 339		
QY	95	CGGTACACCAAGCTGTGTCGCGGACAGCGAGGAGGGCAAAG---CCCCGCGAGAAATCGTAC 151		
Db	340	CGGTACACAAGCTCCTCTCTGGACAGTGAGGACTGCGGGTGGCCACACAGAAATCTCTAC 399		
QY	152	AGTCTCAGCAGACCCCTGAAGGCTTAGACACGAGACGCCCGCTAGCTATGGCAGCGC 211		
Db	400	AGTCTCAGTGAGACTCTGAAGGCGCTATGACCATGACAGCAGGATGCACTATGGAACCGA 459		
QY	212	GTCAAGACCAATTGTCCGCGAGGAGCGGAGAAATTTCTCCGCGACAGGTGCCAATTCACC 271		
Db	460	GTACAGACCTCATCCACCGGAGTCAGATGAGTTTCTTAGACAAAGAAACCACTTCACC 519		
QY	272	CTCGGAGGTGGGGCTGGAAGAAAGTAACGCCGCCCTCACGGGACCCCTGTATCCGGACAGAC 331		
Db	520	CTTGCCGAACTGGGCATCTGTGA---GCCCTCCACACCGGAGCGGCTACTGTCCGAC 576		
QY	332	ATTGGCTCGCCCAATCGCGCTACTCCTATGGGGCTGGCTCTGATGCCACATGGAGGCT 391		
Db	577	ATGGGATCTTTCACAGAGGCTACTCCTCTTAGCAAGGGTCTGAGCGCGACTCCGACACC 636		
QY	392	GACACGGTGTGTCCCTTGACACACCCCGCTGCTGTGGGGCCCGGACACACGGTTCAGGG 451		
Db	637	GAGGAGGGATGTCTCCAGAACACGCCATCAGACTGTGGGCGACAGAGTAATAATCCAGG 696		

Qy	452	CGAGCTCTGCGCTGTCAGCCGGGCGCAATTC	CAATCTCACATCCACCGACACCGAGCAT	511
Db	697	CGAGATTCGGGCTGTGCTCAATCGTGTAAAC	CTCGGCCCTTACCCCTGACTGACTCTGCAAC	756
Qy	512	GAATACTGAGACTGATCATCGGGCGGCC---	TGCAGAACACACGCGGGTTCGGAC	567
Db	757	GAATAAATCAGATGATGAAACGGTCTGTC	CCATTCCACCTACATCTCTCGCTAGTCTC	816
Qy	568	GCGCGCGCGCGCTCTCGCAAGCCACACCC	CAACCCACACGACACCGCGGCTCCATTAA	627
Db	817	CTCCATCTGCTCAGCTGCTAGCTCCCATAT	CTCTCCACCACTAGTGTCC--AGATGCC	875
Qy	628	CTCCCTGAACCGGGGCAATTTCACGCGGAG	AGACACCCAGCGCGCCCGCCCGGACCA	687
Db	876	ATTGTAGACAGCAACACTCTCCCATCAAT	CAATCATGAGACACCAACCTTGAGGAAT	935
Qy	688	CTCGCTCTCGGAGAGCCCCCTCGCGCGG	CGCCAGGAGCTGTCCACCGCCAGAGAA	747
Db	936	CCCCAATTCATCTGCTCAGAGCATGCT	CAGGGCCCGCAGCAAGCTCCAGCAGTGGCC	995
Qy	748	CTGGTGTCTCAACAGCAACATCTCCCT	TGAGACAGCAACCTAGGCAAGCGCATTCCT	807
Db	996	TCCGAACCAACACAGCCAGTCTGAGT	CGCC-----CCTCTCCC	1037
Qy	808	AGGGACATTCGAGCAACACTCATTTGAGAT	TGGACATTCGCGGCGCTCCCGCCATGATGG	867
Db	1038	ACCCCTCAACAACACAGCTGTGCCAT	CACCACTCGTCCGCCAATCTCCCTCAACAGAA	1097
Qy	868	GGTTACAGTACGCGGCACTCTCTTCAAG	CGCTGAGGCACTCCCGGCTCTTGTGCAC	927
Db	1098	CTCACTGACCAATCGCGGAGTCAGAT	TCCAGCGCCCGGCCCAACGACCTGGC	1157
Qy	928	CACATCACAGGATACCACTGACGTCCAG	CACAGTGTACTCTCTCCGCGCCGACCCCT	987
Db	1158	CACACACAGATCTCGTTGAGTTCAG	ACAGCTGGGTCTTAAACAGCAACGTGCCACT	1217
Qy	988	GCCTCCGAGCACCTTCGCGCGCGGCGCT	TTAACCTCAAGAACCCCTCCAAGTACTGTAA	1047
Db	1218	GGAGACCCGCACTTCTCTTCAAGACCT	CTCCGGGAGCACACCT-----T	1265
Qy	1048	CTGGAAGTGGCGAGCCCTGAGCGGCAT	CTGTCATCTCAGCCACTCTGGTCACTGTCTGGC	1107
Db	1266	GTTTACAGAGCTCTTCCCGGGATAC	CCCTTTGACCTCAGGAACGGTTTACACGCCCCGCGC	1325
Qy	1108	ATACTTTGTGGCATGACCTGTTTGGCCT	TAAACTGGCACCTCAGCGCGATGAGGGCA	1167
Db	1326	CCGCTGTCTCCAGGAATACTTTCTCCA	-----GGAGGCTTTCAA	1367
Qy	1168	GATGTAAGATCAAGGAGGACACAGCCAG	ACTTGGCCTGTGCCAACCGACTCTCC-C	1226
Db	1368	GCTGAAGAAGCCCTCCAATACTCTGAG	CTGGAATGTGTCTGCCCTCTCCGCCATTGGCGC	1427
Qy	1227	TATACCCCTCAGGGGCACTGGCTTAGAG	ACCCTTGACAGGAAGCAAGAACACACAG	1286
Db	1428	GGCCCTCTCTTGGCTATTTTGTGGG	GTATTTCTAGTCCCTGGTCTGTGAAACACAG	1487
Qy	1287	AAGGAAGCCAGTAGTTCTTTTCCAGAG	GACAGTTTTCATAGATTTCTGGAGAAATGATG	1346
Db	1488	CAG-----CATAGACAGTGGTGAAG	CAGAAG	1513
Qy	1347	TGGGAAGCGAGCCTCCAGAGATTCCT	CTCGGCACTTCTCGAGATCTCAAGTGTCA	1406
Db	1514	TTGGTCGGCGGGTAAACACAGAAAGT	CCCAACAGCGGGTGTTTTGGAGGTCACAATTCACA	1573
Qy	1407	TAGACCATCTGTGCATCTCAAAATCAAT	TGAATTCAAATGTCTCTGGAAAGGACGCGCTGTTGGCA	1466
Db	1574	TCAGTCAGCCCCAGTCTTAAAGTT	CNAATCTCCCTCGGGAAGACGCTCTCTTTGGTG	1633
Qy	1467	TTTATGGCAAAAGGCGCTCCCTCTTTC	ATCATAACAGTTTGTGAGCTGTGCTGG	1526
Db	1634	TTTACATAACAGAGACTTCCACCATCT	CATGCCAGATGACTTTCATGGAACGTCTGG	1693

Db 3833 CCAGTGCACGGCTTGTGTAAGGCAACAAGCTGCTGGCCCCAGTGGCTCTGGCTGTG 3892
 Qy 3747 GCTCTGACGGAGCCTCTATGTGGTGATTTCACTACATAGNAGATCTTCCCTCTG 3806
 Db 3893 GAATCGATGGAGCCTCTATGTGGTGACTTCAATTACATCCGACGATCTTCCCTCTC 3952
 Qy 3807 GAAATGTCACCAACATCCTAGAGCTGAGGAATAAGATTTTCAACATAGTCAACAGTCCAG 3866
 Db 3953 GAAATGTGACCAAGCATCTTGAGATTACGAATTAAGATTTTAAACATAGCAACACCCAG 4012
 Qy 3867 CACACAAATACTACTCTGCCACAGACCCCTATAGTGGGGCGTCTTCCCTTCTGACAGCA 3926
 Db 4013 CACACAAATGACTACTTGGCAGTGAACCCCGTGTCCGGCTCGCTTACGTTCTCGACACCA 4072
 Qy 3927 ACAGCCGGCGGTCTTTAAATCAAGTCCACTGTGGTGGTGAAGAGCTTTGTCAAGAACT 3986
 Db 4073 ACACAGAGAGNATCTACCGCTCAAGTCTCTGAGTGGAAACCAAGAGCTCGCTGGGAAT 4132
 Qy 3987 CTGAGGTGGTTCGGGACAGGTGACCAAGTCCCTCCCTTTGATGACACTCGCTGGGGG 4046
 Db 4133 CGGAAGTTGTGGCAGGAGCGGAGAGCAGTGTCTACCCCTTTGATGAAGCCGCTGGGGG 4192
 Qy 4047 ATGGTGGAGGCCCACAGAGCCACATCAACCAATCCAGGGTATTACAGTGGCAAGT 4106
 Db 4193 ATGGAGGAAGGCCATAGTGAACCTGTATGAGCCGAGAGGTATTGCAAGTGAACAAGA 4252
 Qy 4107 TTGGGCTGATCTACTTCTGATGGACACCATGATCAGACGCATCGATCAGAAATGGGATCA 4166
 Db 4253 ATGGGCTCATGTAATTTGTGATGCCACCATGATTCGGAAGTTTGACCAAGATGGATCA 4312
 Qy 4167 TCTCCACCTGCTCGGCTTAATGATCTCATCAGCCGGGCACTCAGCTGTGATCTG 4226
 Db 4313 TCTCCACCTGCTGGGCTCCAATGACCTCACTGCGCTCGGCGCGCTGAGCTGTGATTTCCA 4372
 Qy 4227 TCATGGATATTTCACAGTAAGACTGGAGTGGGCCCAAGACTTACGATCAACCCCAATGG 4286
 Db 4373 GCATGGATGTAGCCAGAGTTGCTCTGGAGTGGCCCAAGACCTTGTCTCAATCCCATGG 4432
 Qy 4287 ACACACTCACTTATGCTCTGACAAACAAATGTGTCTGCAAAATCTTGAATAACCAACAGG 4346
 Db 4433 ATAACTCTCTGTATGTTCTAGAGAAACAATGTCTATCTTCAATCACCGAGAAACCAACAG 4492
 Qy 4347 TGCCCATGTGCGGGAGGCCATGCACCTGCCAGGTCCCTGGCATTTGACCACTTCTG 4406
 Db 4493 TCAGCATCAATTGCGGGAGGCCCATGACCTGCCAGTTCCTGGCATTTGACTACT---CAC 4549
 Qy 4407 TAAGCAAGGTGGCCATCCAGCAACCTCTGGAGTCAGCCACCGTCTTGGCTGTGTTTCAACA 4466
 Db 4550 TCAGCAAACTAGCCATTCACCTCTGCCCTGGAGTCAGCCAGTGGCCATTTGCTTCTACA 4609
 Qy 4467 ATGGGTCCTGTATATTGCTGAGACTGATGAGAAAGATCAACCGCATCAGGACGTCA 4526
 Db 4610 CTGGGGTCTCTACATCACTGAGACAGATGAGAAAGATTAACCGTCTACGCCAGGTAA 4669
 Qy 4527 CCAGTAGGAGAGATCTCACTCGTGTGGGGCCCCAGTGGGCTGTGACTGTAAATAATG 4586
 Db 4670 CAACCAACGGGAGATCTGCCCTTTAGCTGGGCGAGCCTCGGACTGGCACTGCAAAAACG 4729
 Qy 4587 ATGCCAATGTGATTTTCTGGAGACGATGTTATGCAAGGATGCAAAATTAATA 4646
 Db 4730 ATGTCAATTGCACTGTTTTCAGGAGATGATGCTACGGGACTGATGCCATCTTGAATT 4789
 Qy 4647 CCCCATCTTCTGCTGTGTGCTGATGGGAGCTTACGTGGCGGACCTTGGGAACA 4706
 Db 4790 CCCCATCATCTTAGCTGTAGCTCCAGATGTTACCATTTACATTGCAGACTTGGAAATA 4849
 Qy 4707 TCCGAATTCGGTTTATCCGGAAGAACAAAGCCTTTTCTCAACCCAGAACATGATGAGC 4766
 Db 4850 TTGGGATCAGGGGGTTCAGCAAGAACAAAGCCTGTTCTTAATGCTTCAACAGTATGAGG 4909
 Qy 4767 TGTCTTACCAATTGACAGAGCTCTATCTGTTTGTATCCACGGCAAGCAGCTGTACA 4826
 Db 4910 CTGCATCCCCCGGAGAGAGGATTAATGTTTTCAAACGCTGATGGCATCCACCAATACA 4969

Qy 4827 CCCAAGCTGCCACAGAGACTACCTGTACAACTTCCACCTACACTGGGACGGGACA 4886
 Db 4970 CTGTGAGCCTGTGTACAGGGAGTACTTGTACAACTTTCATATAGTACTGCAATGATG 5029
 Qy 4887 TCACACTCATCAGACACAACTATGCAACATGGTAAATGTCCGCGAGACTCTTACTGGGA 4946
 Db 5030 TCACATGAATTGATTGACAAATATGGAATTCCTTGAAGATCGTCCGGACAGCAGTGGCA 5089
 Qy 4947 TGGCCCTCTGGTGTGTGTCACAGATGGCAGGTGTACTGGGTGACCATGGGACCAACA 5006
 Db 5090 TGGCCCGTCACTGTGCTGATGCTGACCAACAGATCATCACCTCACCCTGGGACCAAAATG 5149
 Qy 5007 GTGCACTCAAGAGTGTGACCAACAGACACAGTGTGCCATGATGACATACCATGGCA 5066
 Db 5150 GAGGCTCAAAGTCTGTGTCCACACAGAACCTGGAGCTTGGTCTCATGACCTATGATGCA 5209
 Qy 5067 ATTCCGCGCTTCTGGCAACCAAAAGCAATGAAAACGATGAGACACATTTTATGATGAG 5126
 Db 5210 ACATCGGGTCTCTGGCCCAAGAGCGATGAAACAGAGTGGACGACTTTCTTATGACTATG 5269
 Qy 5127 ACAGCTTTGGCGCTGTCAAAATGTGCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAA 5186
 Db 5270 ACACAGAGCGCTGTGACCAACGTTGACGGCCCAACGGGGTGTAAACAGTCTGACCC 5329
 Qy 5187 GTGATACAGACAGTTTCAAGTGTGATGCTCCAGGTAGAGACCTTCA---GCAAGGATGATGA 5243
 Db 5330 GGGAAATGGAGAAATCTATTACCAATTGACATTTGAGAACTCCAAACCGTATGATGACGTCA 5389
 Qy 5244 CCATTAACCAACCACTGTCTGCTCAGCGCTTCTACACACTGTGCAAGACCAAGTCC 5303
 Db 5390 CTGTATCAACCACTCTCTTCTCAGTAGAGGCTCTCTACACAGTGTCAAGATCAAGTTTC 5449
 Qy 5304 GGAACAGCTACTACATCGGGGCGGATGGCTCTTCCGGCTGTCTGCGCCCAACCGCATGG 5363
 Db 5450 GGAACAGCTACTACAGCTCTGTAATAATGTTACCTTGAAGGTGATGCTATATGGGATGG 5509
 Qy 5364 AGGTGGCGCTGAGACTGAGCCCCACTTGTGGCTGGCACCGTCAACCCACCGTGGGCA 5423
 Db 5510 GTATCAGCTTCCACAGCGAGCCCCATGCTCTAGCGGGCACCATACCCCCACCATTTGAC 5569
 Qy 5424 AGAGAAATGTCAGCTGCCATCGCAACAGGGCTCAACTGTGTGAGTGGCGCCAGCGCA 5483
 Db 5570 GCTGCAACATCTCTGCTGCTATGGAGATGGCTTAAATCTCAATTGAGTGGCGCTTAAGNA 5629
 Qy 5484 AAGACAGGCTCGGGGCCAGGTCACTGTCTTTTGGCGCGGCTGCGGGTGCACAAACCGAA 5543
 Db 5630 AGGACAGATTAAAGCAAAAGTCAACATCTTTGGCAGGAAGCTCCGGGTCCATGGAAGAA 5689
 Qy 5544 ATCTCTTCTCTGAGCTTGTATCGGTAAACAGCACAGAGAGATCTATGATGACCCACC 5603
 Db 5690 ATCTCTTGTCCATTGACTATGATCGAAATATTCGAGCTGAAAGATCTATGATGACCCACC 5749
 Qy 5604 GCAAGTTCAACCTTCCGATTTCTGTACGACACAGCGGGGCGGCCAGCCTCTGTCACCCA 5663
 Db 5750 GGAAGTTCACTGAGGATCAATTTATGACCAAGTGGCGCGCCCTTCTCTGCTGCCCA 5809
 Qy 5664 GAGCAGGCTGAATGGTGTCAACGTGACATCTCCCTCGGGGGTTACATTGCTGECATCC 5723
 Db 5810 GAGCGGGCTGGCAGCTGTCAACGTGTCTACTTCTTCAATGGCGCGCTGCGGTGGGCTTC 5869
 Qy 5724 AGAGGGCATCATGTCTGAAAGATGGAATACGACAGCGGGCGCGCATCACATCCAGGA 5783
 Db 5870 AGCGTGGGCCATGACGAGGAGAGACATCGAACGAGCCCGCATCTGTTGCCGCA 5929
 Qy 5784 TCTTCGCTGATGGGAAGACATGGAGCTACACATATTAGAGAGTCCATGTTGCTGCTAC 5843
 Db 5930 TGTTCGCTGACGGGAAGTGTGGAGCTACTCTCTACTTGTGACAAAGTCCATGTTCTCTGC 5989
 Qy 5844 TACACAGCAGAGGAGTATATCTTTGAGTTCGACAAAGATGACCGCTCTCTTCTGTGA 5903
 Db 5990 TTACAGCCCAAGCTCAGTATATATTGATGATGATCTCTCTGACCGCTCTCTTGGCGTCA 6049

Db 8189 GCATCGCTATGGCTCACCCCGACACCCTGGAGAGAGAGAGCCCGGCTCTGTGACC 8248
QY 8112 TGGCCCGGCAGAGCCGTGGCCCAAGCGTGGCCCGGCAGCAGAGACTCGGGAG 8171
Db 8249 AGGCGAGACAGAGGCGCTGGGACGCGCTGGGCAAGAGGAGCAGAAAGCAGGGACG 8308
QY 8172 GGGAGGAAGGCTCGGCGCTTGACAGAGGGGAGAGAGCAGAGTGCTGAGCACAGGGC 8231
Db 8309 GGAGAGGGGAGCGCGCTGTGACTGAGGGGAGAGAGCAGAGCTTCTGAGCACCGGC 8368
QY 8232 GGGTGAAGGCTAGCAGCGCTTTTCTGATCTCTGTCGAGAGTACCCAGAACTGTGAC 8291
Db 8369 GCGTGAAGGCTAGCAGGGATATTACGTGCTTCCGTTGGAGCAATACCCAGAGCTTGCAG 8428
QY 8292 ACAGCGCCAAACAATCCACTTCATGAGACAGAGCGAGATGGCGGAGGTGACAGA 8348
Db 8429 ACAGTGAAGCAATCCAGTTTAAAGACAGATGAGATGGGAAGAGATACAAA 8485

RESULT 12

US-09-800-198-7
; Sequence 7, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herimann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 9826
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (280)..(8478)
US-09-800-198-7

Query Match 40.5%; Score 3385; DB 10; Length 9826;
Best Local Similarity 64.7%; Pred. No. 0; Mismatches 2790; Indels 154; Gaps 18;
Matches 5393; Conservative 0;

QY 35 ATGACGTGAAGGAGAGAGAGCTTTACCGTCTGCTGACCGCGCGCGCGCGAGCGC 94
Db 280 ATGATGTAAGGACCGCGACACCGCTTTTGACAGAGAGACGCTGTGCAAGAGTGT 339
QY 95 CGCTACACAGCTGCTCGGCGGACAGCAGAGAGGGAAG---CCCGGAGAAATGCTAC 151
Db 340 CGCTACACAGCTCCTCTGACAGTGAAGACTGCGGCTGCCACACAGAAATCCTAC 399
QY 152 AGCTCCAGCAGACCTTGAGGCGCTACGACAGAGCGCGCGCTAGCTATGCGAGCGC 211
Db 400 AGCTCCAGTGAATCTGAGGCGCTATGACATGACAGAGAGTGAATGAAACCGA 459
QY 212 GTCAGAGACATTTGCGCGAGAGGCGGAGAAATTTGCGCGCAGAGTGCCAACTTCACC 271
Db 460 GTCACAGACCTCATCCACCGGAGTCAGATGAGTTTCTAGACAAAGAACTTCACC 519
QY 272 CTGCGGAGCTGGGCTGGAAGAGTAACGCCCTCAGCGGACCTGTACCGGACAGAC 331
Db 520 CTTCGCGAACTGGGCACTGTGTA---GCCCTCCCCACACCGAAGCGGCTACTGCTCGAC 576

QY 332 ATTGGCTGCCCCAATATCGGCTACTCCATGGGGGTGGCTCTGTATGCGCAGATGAGGCT 391
Db 577 ATGGGATCTTTCACAGGGCTACTCTCTTAGCAGAGGTCTGACCGCGACTCCGACACC 636
QY 392 GACACGGTGTGTCCCTTGAGCACCCCTGTGTGGGGCCGAGCACAGGTGAGGG 451
Db 637 GAGGAGGGATGTCTCCAGAACACGCCATCAGACTGTGGGGCAGAGGATAAATCCAGG 696
QY 452 CGCAGCTCTGCTGCTGCGCGGGCCCAATTCCAATCTCACACTCACCGACACCGAGAT 511
Db 697 CCAGTTCGGGCTGTGAGTGTGAATCTGGCCCTTACCCCTGACTGACTCTGACAAAC 756
QY 512 GAAACACTGAGACTGATCATCCGGGCGGCC---TGACAGAACCAACGGCGGCTCCGGAC 567
Db 757 GAAACAAATCAGATGATGAGAACGCTGCTCCCAATTCACCTACATCTCGCTAGTCTC 816
QY 568 GCGCGCGCGCGCTCTGCGAGCCCAACACCCCAACAGCAGCAGCGGCTCCATTAA 827
Db 817 CTCCCATCTGCTCAGCTGCTAGCTCCCATTAATCTCCACCATGTAGTGCC-AGATGCC 875
QY 628 CTCCCTGAAACCGGGCAACTTCACGCCGAGGAGCAACCCAGCCCGGCCCCCACGACA 687
Db 876 ATTGCTAGACAGACACCTCCCATCAATCATGACACCAACCTGATGAGAAATCTC 935
QY 688 CTCGCTCTCGGAGAGCCCTCTGCGGGCGGCCCGAGGAGCTGCCAGCCCGCAGAGAA 747
Db 936 CCCCAATTCATACCTGCTCAGAGCATGCTCAGGGCCCGCAAGGCTCCAGCAGTGGCC 995
QY 748 CTGGTGTCTCAACAGCAACATCCCTCTGAGACAGAGAACCTTAGSCAAGCAGCCATTCT 807
Db 996 TCGGAACACACAGCCAGTGTGACTCTGAGGCC---CCTCTCC 1037
QY 808 AGGGAATTTGAGCAACCTCAATTTGATGAGACTTCGCGGCTCTCCGCGCATGATGG 867
Db 1038 ACCCTCTCAACACCAACAGCTGTCCATCAACCACTCGTCGCGCAACTCCCTCAACAGAA 1097
QY 868 GCTTACAGTGAAGGCGCTCTCTTCAAGCTGGAGGCACTCCCGCTCTTCTGAC 927
Db 1098 CTCCTGACCAATCGCGGAGTCAAGTCAACGCCCGCGCGCGCGCAATGACTGGC 1157
QY 928 CACATCACAGGGTATCCCACTGACGCTCCAGCACAGTGTACTCTCTCGCGCCGACCCCT 987
Db 1158 CACACACAGAGTCCGTTTCACTTCAGACAGCTGGTGTAAACAGCAACGTCCTACT 1217
QY 988 GCCCGCAGACCTTTCGCGCGCGCTTTAACTCAAGAGCCCTTCAAGTACTGTAA 1047
Db 1218 GGAGACCCCGCACTTCTCTTCAAGACCTCTCGGGGAGCACACCT-----T 1265
QY 1048 CTGGAAGTGGCAGCGCCCTGAGCGCATCGTCACTCTCAGCACTCTGCTCATCTGCTGC 1107
Db 1266 GTTACGAGCTCTTCCCGGGATACCTTTGACCTCAGAACGTTTACCGCCCGCC 1325
QY 1108 ATACTTTGGCCATGCACTTTTGGCTTAACTGGCACTTGACCGCATGAGAGGGA 1167
Db 1326 CGCGCTGTGCGCAGAACTCTTCTCCA-----GGAAGCTTTCAA 1367
QY 1168 GATGTATGATCAGGAGGACAGCCAGAGTTGGCTGTGCAACCGAGCTCC-C 1226
Db 1368 GCTGAAGAAGCCCTCAAAATCTCAGCTGGAATGTGTGCGCTCTCGCCATTTGCCGC 1427
QY 1227 TATACCTCTCAGGGGCACTGGCTTAGAGACCCCTCAGAGGAAAGGCAAGAACACAG 1286
Db 1428 GCGCTCTCTTGGCTATTTTGTGCGCTATTTTCATAGTGCCTGTGCTTGAAGACAG 1487
QY 1287 AAGGAAAGCCAGTAGTATTTCTTCCAGAGGACAGTTTCTAGATTTCTGAGAAATGATG 1346
Db 1488 CAG-----CATAGACAGTGGTGAAGCAGAG 1513
QY 1347 TCGGAAGGCGAGCTCCAGAGATTTCTCTGCGCACTTTCTGGAGATCTCAAGTGTCA 1406
Db 1514 TTGCTGCGCGGTATACAGAGTCCACAGGGGTGTTTGGAGGTTCACAAATTCACA 1573
QY 1407 TAGACCATCTCTGCTATCTGAATTTCAATGTGTCTCTGGGAAAGGACGCCCTGTGGTGA 1466

Db 5930 TGTTCGCTGACGGGAAGTGTGAGCTACTCTTACCTTGACAGTCCATGGTCTCTCTGC 5989
Qy 5844 TACAGCCAGAGCAGTATATCTTTGAGTTGCACAAAGATACCSCCTCTCTTTGTGA 5903
Db 5990 TTCAGAGCCAAAGTTCAGTATATATTTGAGTATGACTCTCTGACCCCTCTCTGCGTCA 6049
Qy 5904 CGATGCCCAAGCTGGCGGCAGACACATAGAGACCATCCGCTCAGTGGGCTACTACAGAA 5963
Db 6050 CCATGCCCAGGTGGCCCGGCAGAGATGTCCACACACCTCCATCGGCTACATCCGTA 6109
Qy 5964 ACATCTATCAGCCCTTGAGGGCAATGCCCTCAGTCATACAGGACTTTCACGTGAGGATGGCC 6023
Db 6110 ATATTTACACCCGCTGAAAGCAATGCTTCGCTCATCTTTGACTACAGTATGACGGCC 6169
Qy 6024 ACCTCTTCACCTTCTACTGGGCACTGGCCGAGGGTGATATACAAAGTATGGCAAC 6083
Db 6170 GCATCTGAAGACCTCTCTTTTGGGACCGGAGCCAGGTGTTCTTCAAGTATGGGAAC 6229
Qy 6084 TGTCAAAGCTGGCAGAGACGCTCTATGACACCAACCAAGGTCAGTTTCACTATGACGAG 6143
Db 6230 TCTCCAGTTATCAGAGATTGTCTACGACAGTACCGCGCTCACCTTCGGGTATGACGAG 6289
Qy 6144 CGGAGGCATGTGAAGACCATCAACCTACAGAAATGAGGCTTTCACCTGACCATCCGT 6203
Db 6290 CCATGTGTCTTGAAGATGTCACCTCCAAAGTGGGGCTTCTCTGCAACCATCAGT 6349
Qy 6204 ACCGTAGATTGGCCCTCATTTGACGACAGATCTTCCGCTTCACTGAGGAGGATGG 6263
Db 6350 ACCGAAGATTGGCCCTCGTGGCAAGAGATCTACAGTTTCTCCGAGGAGGATGG 6409
Qy 6264 TCACGCCGCTTTGACTACAACT---ATGACAAACAGCTTCCGGGTGACCAAGCATGACGG 6320
Db 6410 TCAATGCCAGTTTGACTACACCTATCATGACAAAGCTTCCGCTCGCAAGCATCAAGC 6469
Qy 6321 CTGTGATCAGAGACCCCACTGCCATTCATCTCTATGCTATGATGATGTTCAGGCA 6380
Db 6470 CCGTCAATGAGATCCCTCCCGTTGACCTCTACCGCTATGATGATTTCTGGCA 6529
Qy 6381 AGACAGAGCTTTGGGAAGTTTGGTGTCACTTATGATGACATTAACCAAGATCATCACC 6440
Db 6530 AGGTGAACACTTTGGTAACTTTGGAGTCACTATTTATGACATCAACCAAGATCATCACC 6589
Qy 6441 CAGTGTGATGACCCACCAAGACATTTGATGATATGCGAGATGAAGGATGTCAGT 6500
Db 6590 CTGCGGTGATGACCTCAGAAACACTTCGACACCATCGCGGATCAAGGAGGTCCAGT 6649
Qy 6501 ATGAGATCTTCCGCTCGCTCATGTCTGATGACCGTCCAGTATGATGATGATGATGATG 6560
Db 6650 ATGAGATGTTCCGCTCCCTCATGCTGATGATGATGATGATGATGATGATGATGATGATG 6709
Qy 6561 TAGTGAAGAGAGCTGAAGGTAGGACCCCTAGCCCAATACCACTCGCTACTCTATGAGT 6620
Db 6710 TGATCAAGAGGAGCTAAAACCTGGGCGCTTATGCCAATACCAAGAGTACACCTATGACT 6769
Qy 6621 ATGATGCTGACGCGCAGCTGCACAGCTCTCCATCATGATGACAGCCACTCTCGCGCTACA 6680
Db 6770 ACATGGGAGCGGAGCTCCAGAGCTGGCGCTCAATGACCGCCGACCTGGCGCTACA 6829
Qy 6681 GCTACGACCTCAATGGGAACCTGCACTTATGAGCCCTGGGAACAGTGCACGCTCAGCTCAC 6740
Db 6830 GCTATGACCTTAATGGAACTCTCCACTTATGAAACCCAGGCAACAGTGTGCGCTCATGCT 6889
Qy 6741 CACTAGGATGATACATCCGAGCGGATCACTCGGCTGGTGCAGTGCAGTGCAGTGCAGTGC 6800
Db 6890 CCTTGGCTATGACCTCCGGATTCGGATACCAAGCTCGGGGATGTGAGTACAAATG 6949
Qy 6801 ATGAGGATGCTTCTGAGGACGCGGGGGTGTATCTTTGAGTACAACTCAGCTTGGCC 6860
Db 6950 ACCAGATGCTATCTGTGCCAGAGAGGTCTGACATCTTCGAATACATTCGAAGGCC 7009
Qy 6861 TGTCTATCAAGGCTACACCGGGCTGGCAGCTGGAGTGTGAGTACCGCTACGATGGCC 6920

Db 7010 TCCTAAAGAGACCTACAAAGCCAGCGGTTGAGTGTCCAGTACCGCTATGATGGCG 7069
Qy 6921 TGGGGCGGCGGTGTCCAGCAAGAGAGCCACAGCCACCACTTGCACTTCTTCTATGCGAG 6980
Db 7070 TAGGACGCGGGGCTTCTTACAAAGACCAACCTGGGCCCACTGCGAGTACTTCTACTCTG 7129
Qy 6981 ACCTGACCAACCCCAACCAAGGTACCCACCTGTACAAACCACTCCAGCTCTGAGATCACCT 7040
Db 7130 ACCTCCACACCCAGCGGATCACCCATGTCTACATCACTCCAACTCCGAGATTACCT 7189
Qy 7041 CCCTCTACTAGCTTCAAGGACACCTCTTTGCAATGAGAGTACAGTGGTGTGATGAGT 7100
Db 7190 CACTGTACTACGACCTCCAGGGCCACCTCTTTGCAATGAGAGCAGCAGTGGGAGGAGT 7249
Qy 7101 TTTACATAGCTGTGACAAACATCGGACCCCTCTGCTGTCTTTAGTGGAAACAGGTTGA 7160
Db 7250 ACTATGTTGCCCTCTGATAACAGGGACTCTCTGGCTGTGTTACAGCATCAACGGCCTCA 7309
Qy 7161 TGATCAAGCAAAATCCTGTACACAGCCTATGGGAGATCTACATGGATACCAACCCCAACT 7220
Db 7310 TGATCAAAAGCTCAGTACAGGCTATGGGAGATTATTATGATCTCCCAACCCCGACT 7369
Qy 7221 TTGATCATCATAGGCTACCATGGTGGCTCTATGATCCACTCACCAGCTTGTCCACA 7280
Db 7370 TCCAGATGGTCAATGGCTTCCATGGGGACTTATGACCCCTGACCAAGCTGGTCCACT 7429
Qy 7281 TGGGCCGCGAGATTATGATGTGCTGCCGAGAGCTGAGCTAGCCACAGACCAAGAGCTGT 7340
Db 7430 TCACTCAGCTGATGATGATGTCTGCGAGAGATGAGACTCCCGAGACTATACCATGT 7489
Qy 7341 GGAAGCACTTTAGTAGCAGCAAGCTCATGCTTTTAACTCTATATGTTCAAAAACAACA 7400
Db 7490 GGAAGAAAGCT---GGGCAAGAGAGCGGCCCTTTTAACTGTATGTTTCAAGAGCAACA 7546
Qy 7401 ACCCATCAGCAACTCCCGAGAGATCAAGTGTCTTATGACAGATGTTTAAAGCTGGCTGC 7460
Db 7547 ATCTCTCAGAGTGTGAGTATGTTGAGAACTACGTGACAGATGTGAAAGCTGGCTTG 7606
Qy 7461 TCACCTTTGGATTCAGCTACACAACTGATCCCTGGTTATCCCAACCAAGACATGGATG 7520
Db 7607 TGATGTTGGATTTGAGCTTAGCAACATCATCTCTGGCTTCCCGAGAGCCAAATGTATT 7666
Qy 7521 CCATGGAACCTCTCAGAGCTCATCCACACAGATGAAGACGAGAGTGGGCAACA 7580
Db 7667 TCGTGCCTCTCCCTATGA-----ATTGTCAGAGAGTCAAGCAAGTGAAGATG 7714
Qy 7581 GCAAGTCTATCTCTCGGGTACAGTGTGAAGTACAGAGAGCTCAAGGCCCTTTGTCACTT 7640
Db 7715 GACAGCTCATTTACAGGTGTCCACAGACACAGAGAGACATACCAGGCCCTTCATGGCTC 7774
Qy 7641 TAGAACGGTTTGACCACTCTATGGCTCCACATCACTCAGCTGCGAGAGGCTCCAAAGA 7700
Db 7775 TGGAA-----GGACAGGTCACTTAAAGAGCTCCACGCGCAGCATCCGAGAGAAAGCAG 7828
Qy 7701 CCAAGAAAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGGGCTCAAGTTTGCCTTGAAG 7760
Db 7829 GTCACTGGTTGGCCACCAACCGCCCATCATTTGGCAAGGATCATGTTTGGCATCAAG 7888
Qy 7761 ATGCGCGAGTGAACACAGACATCATCAGTGTGCGCAATGAGGATGGGCGAAGGTTGCTG 7820
Db 7889 AAGGCGGGGTGACCAACCGGGGTGTCCAGCATTCGACGCAAGATAGCCGCAAGGTGGCAT 7948
Qy 7821 CCATCTTGAACCACTGCGCCACTTACTAGAGAACTTGACATTCACCATTTGATGGGGTGGATA 7880
Db 7949 CTGTGCTGAACACCGCTTACTCTGGAAGATGACATACAGCATCGAGGGCAAGGACA 8008
Qy 7881 CCCATTACTTTGTGAACACAGGACCTTTAGAAGGTGACCTTGGCCATCTCTGGGCTCAGTG 7940
Db 8009 CCCACTACTTTGTGAAGATTGGCTCAGCCGATGGCGACTTGGTCACTAGGCACCAACA 8068
Qy 7941 GGGGGCGCGCAACCTTGAGAAATGGGGTCAACCTCACTGTGCTCCAGATCAACACAGTAC 8000
Db 8069 TCGGCCCAAGGTGCTAGAGAGCGGGGTGAACGTGACCGGTGTCCAGCCCAACGCTGCTGG 8128

```

QY 8001 TTAATGCCAGACTAGACGCTACACAGACATCCAGCTCCAGTACGGGGCACTGTGCTTGA 8060
Db 8129 TCAACGCCAGACTCGAAGGTTCAACAACTTCCAGTTCAGTACTCCAGCTGCTCTCA 8188
QY 8061 ACACACGCTACGG-----GACAACTTTGATGAGGAGAGGACGCGTCTTGAGC 8111
Db 8189 GCATCCGCTATGCTTACCCCGCACACCTGAGCAGAGAGAGGCGCGTCTTGACC 8248
QY 8112 TGGCCCGGACAGAGCGCTGGCGCAAGCTGGGCGCGGAGCAGAGAGACTGCGGGAAG 8171
Db 8249 AGCGGACAGAGAGGCGCTGGGCGCGGCTGGGCGCGGAGGAGCAGAGAGAGGAGCG 8308
QY 8172 GGGAGGAGGCGCTGGGCGGCTGGACAGAGGCGGAGAGAGGAGGCTGAGCAGAGGCG 8231
Db 8309 GGAGAGGAGGCGGCGCTGTGACTGAGGCGGAGAGCAGAGCTTCTGAGCAGCGGCG 8368
QY 8232 GGGTGAAGGCTACGAGCGCTTTTTCGTGATCTCTGTGAGAGAGTACCCAGAACTGTAG 8291
Db 8369 GCGTGAAGGCTACGAGGAGTATTACGTGCTTCCCGTGGAGCAATACCCAGAGCTTGCAG 8428
QY 8292 ACAGCGCAACACATCCACTTCTATGACAGAGCGAGATGGCGGAGGTGACAGA 8348
Db 8429 ACAGTAGCAACATCCAGTCTTTTAAAGCAGATGAGATGGGAAAGAGTAACAAA 8485

RESULT 13
US-09-808-602-74
; Sequence 74, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Heriman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115a1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-808-602-74

```

```

Query Match 40.58; Score 3384.8; DB 9; Length 8797;
Best Local Similarity 64.8; Pred. No. 0;
Matches 5467; Conservative 0; Mismatches 2712; Indels 255; Gaps 20;

QY 35 ATGACGCTGAAGGAGAGAAAGCTTACCGCTCGCTGACCGCGCGGCGCGAGCGGCGC 94
Db 113 ATGATGTAAAGGACCGGCGACATCGCTCTTTGACAGGAGCGGTGTGCAAGAGTGT 172
QY 95 CGCTACACAGCTCGTCGGGGACAGAGAGGAGGAAAG---CCCCGAGAAATCGTAC 151
Db 173 CGCTACACAGCTCGCTCTGACAGTGAAGGACTGCGGCTGTGCGGCACTCAGAACTCTAC 232
QY 152 AGCTCCAGCGAGACCTCGAGGCGCTACGACAGAGCGCGCGCTAGGCTATGCGAGCGCG 211
Db 233 AGTTCCAGTGAAGCTTGAAGGCTTATGACCAATGACAGAGAGTGAAGTGAAGACGA 292
QY 212 GTCAGGACATTTGCGCGAGGAGGCGGAGAAATTTGCGCGCAGAGTGCCAACTTCAAC 271

```

```

Db 293 GTCACAGACCTGGTGTACCGGGAGTCCGATGAGTTTTCTAGACAAAGGACAAACTTCACC 352
QY 272 CTGCGGGAGCTGGGGCTGGAAGAGTAACGCCCTCACGGGACCTTGATCGGAGAC 331
Db 353 CTGCGAGAAATGGGAATCTGCGAG---CCCTCCCGACACCGAAGTGGTTACTGTGTCGAC 409
QY 332 ATTGCGCTTGCCTCAATATGCGGCTACTCCATGCGGGGCTGGCTCTGATGCCGACATGAGGCT 391
Db 410 ATGGGATCTCTCCACCGGGCTACTCCCTGAGCACTGGGTCTGATGACAGCTCGGACACC 469
QY 392 GACAGGTGCTGTCTCCCTGAGCACCCTGCTGTGCGGCGGAGCAGACAGGTGAGGG 451
Db 470 GAGGAGGAGGATGTCTCCAGAAATGCCATCAGATGTGGGAGCGGAGTAAATTCAGG 529
QY 452 GCGAGCTCTGCTGTCTCCAGCGGCGCAATTTCCAACTCTCACTCACCGACACCGAGCAT 511
Db 530 GCGAGCTCTGCTGTCTCCAGCGGAGCACTCGGCGCTTACTCTGACTGACTCTGACAA 589
QY 512 GAAACACTGAGACTGATCATCCGCGGCGGCG---TGACAGAACACGCGCGGCTCCGGAC 567
Db 590 GAAATAAATGGGATGAGCAATGGTCTGCTCCCAATTCACCTACATCTCTGCTAGCCTC 649
QY 568 GCGCGCGCGCGCTCTCGACGCGCCACACCCCAACCCAGACACCGCGGCT----- 620
Db 650 CTCCCATCTGCTAGCTGCTAGCTCCCATTAATCTCTCCAGGTAGTGTGACAGATGCCA 709
QY 621 ----- 620
Db 710 TTGTAGACAGCAACACCTCCATCAGATCATGACACCAACCCCTGATGAGAAATTCCTC 769
QY 621 -CCATTAAGTCTGTAACCGGGGCACTTACGCGGAGGAGCAACCCCGAGCGCGCGCC 679
Db 770 CCCAATTCATACCTGCTCGGCGCACTCTCGGCGGCGGCGGCGGCGGCGGCGGCGG 829
QY 680 ACGGACCAAC-----TCGCTCTCGGAGAGAGCGCGGCGGCGGCGGCGGCGGCGG 721
Db 830 CCAAAACACACACAGCCAGTCAACACTGAGGCGCGCTCTGCGACCGCTCATACACACAC 889
QY 722 CAGGAGCTGCGGCGGCGGAGAACTGCGTCTTAACAGAGCAATCCCTCTGAGAGACC 781
Db 890 CTGTCCCAACCACTCTCTCGGCGCACTCTCTCAACAGAGAACTCACTGACCAATCGGCG 949
QY 782 AGGAACTTAGCAAGCAGCCATTCCTAGGACATTTGAGGAGCAACCTC-----ATT 832
Db 950 AGTCAATTCAGCGCGCGGAGCTCTCGGCGGCGGAGAGCTTGGCGACACCGCGAGAGTCTGT 1009
QY 833 GAGATGAGCAATTCGCGGCGCTCCCGCCATGATGGGCTTACAGTACGCGGAGCTCTCTC 892
Db 1010 CAGCTCCAGGATAGCTGGGTGCTGAACAGTAACGTCCTCACTGAGAGACTCGGCACTTCTT 1069
QY 893 TTCAA---GCCTGAGGAGCACTCCCGCTCTTCTGCAACATCACTACAGGAGTACCACTG 949
Db 1070 TTCAAACGCTCGTCTGSAAGCAGACCCCTGTTTACAGAGCTCTTCTCGGAGTACCTTTG 1129
QY 950 AGCTCCAGCAGAGTGTACTCTCTCGGCGCGGAGCGCTTCCCGCGGAGCACTTCCCGCGG 1009
Db 1130 ACTCAGGAGCGTTTATACACCAACCGCGCTGCTGCGGAGAAATACATTTCTCCAGG 1189
QY 1010 CCGGCTTTAACTCAAGAGCCCTTCAAGTACTTAACTGSAAGTGCAGGCGGCGGCTGAGC 1069
Db 1190 AAGGCGCTTCAAGCTGAAGAAACCTTCAAACTCACTGAGTGGAAATGTGCTGCGCTGCT 1249
QY 1070 GCATCGCTCATCTAGCGCACTCTGCTGATCTGCTGCGCATCTTTGTGGCGCATCTTTGTG 1129
Db 1250 GCATCGCGCGCGCTCTCTTGGCAATTTGCTGGCATATTTCTAGCAATGATCTG 1309
QY 1130 TTGCGCTTAACTGCGCACTGCGCGGATGAGGCGGAGATGTATGAGATCAGGAGAGAC 1189
Db 1310 CTGCGACTCAATTTGGCACTCCAGCGGAGAGTGAACACACTTTTAAACA-----ATGGC 1363
QY 1190 ACAGCGAGAGTGGCTGTGCGCAACCGAGCTCTCTCTATACCTCTAGGGGCGACTGCG 1249
Db 1364 GTAAGGACGCGCTTACAGGAAACGATGATGTGGCAACAGTGTGCGCATCTGAGGAGCAAGTG 1423

```

QY 1250 TTAGAGACCCCTGACAGGAAAGGCAAGGAACACAGAGAAAGCCAGCTAGTTCCTTT 1309
Db 1424 CCCTGGTCATTGAAAA-----1439
QY 1310 CCAGAGACAGTTTCATAGATCTTGAGAAATTTGANTGGAAAGGGAGCCTCCCAAG 1369
Db 1440 -----ACAGCAGATAGACAGTGGCGAAGCAGAAAGTTGGTCGGCGGTGACACAGGAA 1492
QY 1370 ATTCTCTGSCACTTTCTGGAGATCTCAAGTGTTCATAGACCATCTGTGCACTGAAA 1429
Db 1493 GTCCCAACAGGGGTGTTTGGAGTCCAGATTACATCAGTCAGCCTCAATTTCTTAAAG 1552
QY 1430 TTCAATGTGTCTCTGGGAAAGGAGCCCTGGTGGCAATTTATGGCAGAAAGGCCCTCCCT 1489
Db 1553 TTCAACATCTCCCTGGGCAAGGATGCCCTCTCGGTGTCTATATAGGAGAGGACTACCA 1612
QY 1490 CTTTCACATACAGATTGACTTTGGAGCTGTGAGCTGTGATGGCAGAGGCTCTTAACCCAG 1549
Db 1613 CCGTCTCATGCCCAAGTATGACTTATGGAACGCTCTGGATGAA-----AG 1657
QY 1550 GAGCGGGAGCCTAGAGGGAGACCCCGCCAGTCTCGGGAACTGTGCCCCCTCCAGC 1609
Db 1658 GAGAAATGGAGCGTGTGAGTCCGCTCGCCAGGGAACGCCGAGCATCCAGACTCTGGTGCAG 1717
QY 1610 CATGAGACAGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTTGGCTTTTTCAT 1669
Db 1718 AACGAGCTGTGTGTGTCAGTACTTTGGATGTGGGCTGTGGCACCCTGGCTTCTTACAT 1777
QY 1670 GAGGAAAGAGTCAGAAAGTGTTCCTTCTCACCACCTGCCATTGAGTCGGTGGATAAC 1729
Db 1778 GACGGCAAGACAGAGGATGTCTCTTCAACACTGTTGTCTTAGATTCACTGTCAGGAG 1837
QY 1730 TGCCCCAGCACTCTATGCAATGTGACTGTGATCTCTGGACCTGCCACTGCTTCCTG 1789
Db 1838 TGTCCACGGAACGTGTCAGGGAACGGTGAATGCTGTCTGGACTGTGTCACTGTTCCTCCA 1897
QY 1790 GGTTCCTGGCCCGGACTGTGCAGAGCCTCTGCCCCGTCTGTAGCGGAATGCG 1849
Db 1898 GGATTCCTAGTGTGAGACTGTGCTAAAGCTGCTGCCCCGTGACTGTGCAGCGGAATGGA 1957
QY 1850 CAATACATGAAGCAGATCTGTGCCACAGTGGCTGGAAGCGCTGAGTGCAGATGTG 1909
Db 1958 CAGTATTTAAAGGAACGTGCCAGTGTACAGCGGCTGGAAGGTGCAGAGTGTGATGTG 2017
QY 1910 CCACCAACAGTGTATCGATGTGGCTGCAGCAACATGGCACTGTGCATCAGGGCAAC 1969
Db 2018 CCTATGAACCAATGTATCGATCTCTTCTGTGGGGGCCATGGCTCTCTGCAATTGATGGGAAC 2077
QY 1970 TGCACTGCAACCTGGCTACAAGGGGAGAGCTGTGAGGAAGTGGACTGCAATGCAACCC 2029
Db 2078 TGGGTGTGCTGCTGCTGCTACAGGGGAGCACTGTGAGGAAGTGTGATGCTTGTGATCT 2137
QY 2030 ACATGTTACGGCGGGTGTCTCGGTGAGAGGCAATGCCATTTGTTGTGGATGGGA 2089
Db 2138 ACCTGTCTCAGCAATGTTGTCTGTGTAATGAGAGTGTCTATGCAAGCCCGCTGGGT 2197
QY 2090 GGCACCAACTGCGAGACCCCGAGGCCACATGCTTAGACCAAGTGTTCAGGCCAGCAAC 2149
Db 2198 GGTCTAACCTGTGAGCTGGGAGGGTCCAGTGGCCAGACAGTGTAGTGGGATGSCAT 2257
QY 2150 TTCTCCCGACACCGGGCTTTGCAGCTGTGACCCCAAGCTGGACTGGACACGACTGTCT 2209
Db 2258 TACTCTCCTGACTCCGGCTCTGCAGTGTGATCCGAACTGGATGGTCCCGACTGCTCT 2317
QY 2210 ATCGAGATCTGTGTCGCCACTGTGTGGCCATGGGTGTGGTGGGGGACCTTCCCGC 2269
Db 2318 GTTG---TGTGCTCAGTAGACTGTGGCACTCAGCGGCTGTGCATCGGGGAGCCTTCCCGC 2374
QY 2270 TGGAGGATGGCTGGATGGGGGAGGCTGTGCACAGCGGGCTTGCACCCCGGCTGTGCC 2329
Db 2375 TGTGAAGAGGCTGGACAGCGAGCTGTGTGACCAAGCGGTGTGCCACCCCGCTGCATT 2434

QY 2330 GAGCATGGGACCTCGCGCAGCGCAAGTGCAGTGCAGCCCTGCTGGAAATGGCGAACAC 2389
Db 2435 GAGCAGGGGACCTGTAAAGATGGCAATGTGAATGCCGAGAGGCTGGAATGGTGAACAC 2494
QY 2390 TGCACCATCGTCACTATCTCGATAGGGTAGTTAAAGAGGGTTCCTCCCTGGTGTGCAAT 2449
Db 2495 TGCACCATGAT-----GGCTGCCCTGATTTGTGCAAC 2527
QY 2450 GGCACGCGCAGATGTACCTTAGAAGTGTGGTGGCACTCGCTGCCAGCTGGCTGG 2509
Db 2528 GGTAAAGGAGATGCACACTGGGTGAGAACAGCTGGCAGTGTGTCTCCAGACCGGCTGG 2597
QY 2510 AGAGGAGTGGCTGTGACACTTCCATGGAGACTCCTCGGTGACAGCAAGACAATGAT 2569
Db 2588 AGAGGCTGTGATGCAACGTTGCCATGGAACCTCTCGGTGATGAACAAGGATAATGAG 2647
QY 2570 GGAGTGGCTGTGGAGTGCATGAGCCCTGACTGTGCTCCAGCCCTGTGCCCATATC 2629
Db 2648 GGAGTGGCTGTGGAGTGCCTGGACCCCTGACTGTGCTCTACAGTCAGCCTGTGAGAAC 2707
QY 2630 AACCGCTGTGCTTGGCTCCCTTAAGCCCTCTGACATCATCCAGGAGACACAGGTCCT 2689
Db 2708 AGCTGTCTCTCGCGGGGTCTCGGGAACCCCTTGGACATCATTCAGCAAGTTCAGACAG 2767
QY 2690 GTGTACAGAGAACCTTACACTCTCTATGACCGCATCAAGTTCCTCTGCTGGGAGGAG 2749
Db 2768 TGGCTGTGAGTGA-----AGTCTCTTCTATGACCGCATCAAGCTCTTGGCAGGCAAGGAC 2821
QY 2750 AGCAGCACAATAATCCCGGGGAGAACCCCTTTCATGAGGGGATGCTTGTATTTCGT 2809
Db 2822 AGCACCACATCATCTCTGGAGACACCCCTTCAATAGACGCTGGTGTCTGTATCCGA 2881
QY 2810 GGCCAAAGTATGACATCAGATGGAAACCCCTCGTGTGGTGTGAACATCAAGTTTGTCAAT 2869
Db 2882 GGCCAAAGTATGACATCAGATGGAGTGGACTCTCTTGGTGGGTGTGAATGTCTTTTGTCAAG 2941
QY 2870 AACCTCTCTTTGGATATACATCAGCAGGACAGATGCGAGCTTGGACTTGGTGAACAT 2929
Db 2942 TACCCAAATATGCTACACCATCACTCGCCAGGATGGCAGTTTGACCTGATTTGCCAAT 3001
QY 2930 GCGGCATCTCCATCATCTCGGTTCGAGGGGACCTTTTCATCACACAGGAGCACACC 2989
Db 3002 GGGGTTCTGCTTGTACTCTTTCATTTGAGCGAGCCCTTTTCATGAGCCAGGAGCGACA 3061
QY 2990 CTGTGGCTGCCATGGATCGCTTCTTGTATGGAACCATCATCATGACACATGAGGAG 3049
Db 3062 GTGTGGCTGCCATGGAACAGCTTCTATGCCATGGACACCTTGGTGAATGAAGCCGAGGAA 3121
QY 3050 AATGAGATTCCAGCTGTGACCTGAGCAATTTTTCGCGCCCAACCCAGTCTCTCCA 3109
Db 3122 AACTCCATCCCGACTGTGACTGCTGCTGCTTGTTCGCGCAGATCCATCATCATCTCC 3181
QY 3110 TCCCACTGAGCTCTTTCGCGAGCTCTGTGACAGAAAGGCCCATTTGTGCGGAAAT 3169
Db 3182 TCTCTCTGTCCACTCTTTCAGCGCTTCCCTGCTCGAAACCCCAATTTGCTGCTGAGACC 3241
QY 3170 CAGCTTTGAGGAGGAAATCTCTATCTGCTGGCTGCAAGATGAGGCTGACCTGAGC 3229
Db 3242 CAGTTCTTATGAAGAAATTTGAGTCCCTGTTACCAATGTGAAGTCTCCCTTATCTCAGC 3301
QY 3230 AGCCGACCCCTGGCTACAAATCTGCTCAGGATGACCTCCACCCACCCAGCATCTCC 3289
Db 3302 TCTAGAACTGCAGGGTATAAGTCTGCTGAAGATCAACCATGACGAGTCCACAGTCCCC 3361
QY 3290 TTTCACTCATGAGTGCACCTCATGTTAGCGGTGAGGGCGCTCTTTCAGGAAGTGG 3349
Db 3362 TTGAACCTCATCAGGGTTCACCTGATGTTGCTGTAGAGGGGATCTCTTTCAGGAAGTCA 3421
QY 3350 TTCCCTGCAGCCCGCAGACCTGTCTTATTTTTCATTTGGACAAAGACAGACGCTCTACAAC 3409
Db 3422 TTCCAGGCTTCTCCCAACCTAGCTACACATTCATCTGGGACAAAGACAGATGCTTATGGC 3481
QY 3410 CAGAAGGTGTTGGGCTTTTCAGAGCCCTTGTTCGTGGGTATGAAATGATGATCTCTGC 3469

Db 3482 CAAAGGGTTATGGCCCTATCGGATGCTGTTGTTGTTGGGTTTGAATATGAGACCTGC 3541
 Qy 3470 CCAGATCTAATCTCTGTGGGAAAAAGAAACAAACAGTGTGTCAGGGCTATGAAATTTGACGGC 3529
 Db 3542 CCCAGTCTATCTGTGGGAGAAAAGACAGCCCTGCTTCAGGGAATTCGAGCTGGACCT 3601
 Qy 3530 TCCAGCTTGGAGATGGAGCCTTAGACAAACATCATGCCCCCTCAACATTCAGAGTGTATC 3589
 Db 3602 TCCAACTTGGAGCTGTGCTTGGGAAACAGCAACACCCCTCAATGTGAAAGCGGAATA 3661
 Qy 3590 CTGCACAAAGGGAATGGGAGAACCAAGTTGTGTCAGCAGCCTCTCTGTCATTTGGGAGC 3649
 Db 3662 CTACACAAAGGACAGGGAGAACCAAGTTCTGACCCAGCAGCCTGCTCATCATCAGAGC 3721
 Qy 3650 ATCATGGGCAATGGGCGCCGAGAGAGATCTCTGCCCCAGCTGCAACGCGCTTGTGAC 3709
 Db 3722 ATCATGGGCAACGCTCGCCGAGAGCATCTCTGTCCTGTCAGCTGCAATGGCTTGTGAA 3781
 Qy 3710 GGCACAGCTCTGGGCGCCAGTGGCCCTCACCTGTGGCTCTACGCGGAGCTCTATG 3769
 Db 3782 GGCACAAACTGTTAGCCCTGTGGCCCTGGCTGTGGGATCGATGGAGCCTCTTTGTT 3841
 Qy 3770 GGTGATTTCAACTACATTAGAAGGATCTTCCCTCTGGAAATGTCAACCAATCTCTAGAG 3829
 Db 3842 GGTGACTTCAACTATATCCGGCGCATCTTCCCTCTCGAAATGTGACCAAGTATCTTTGGAG 3901
 Qy 3830 CTGAGGATTAAGATTTACACATAGTCAAGTCCAGCAGACAAATACTACCTGGCCACA 3889
 Db 3902 TTACGAAATTAAGATTTAAACATAGCAACAGCCGAGGACACAACTACTTCTGGCTGTG 3961
 Qy 3890 GACCCCATGATGGGCGCTCTTCTTCTGACGAAACAGCCGGGGTCTTTAAATC 3949
 Db 3962 GACCCGTGACTGGCTCACTTACCTAGCTCTGACACCAACAGTCCCGAATCTACCGAGTC 4021
 Qy 3950 AAGTCCACTGTGGTGTGAAGGACCTTGTCAAGAACTCTGAGTGGTGTGGGGGACAGGT 4009
 Db 4022 AAGTCTCTGAGCGAGGCAAGACCTGGCTGGAATTCGGAAGTTGTGGCAGGGAAGTGGC 4081
 Qy 4010 GACCAGTCCCTCCCTTTGATGACATCGCTGGGGATGTTGGGAGGCCACAGAGCC 4069
 Db 4082 GAACATATGTTACCTTTGATGAAGCCGCTGTGGGATGAGGAGGAGGTGTGGAGCC 4141
 Qy 4070 ACATCCAAATCCAGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTCTGTGGAT 4129
 Db 4142 ACCCTGATAGCCCCAGAGTATTGCAAGTAGACAAGAAATGGGCTTATGTTGTTGAT 4201
 Qy 4130 GGCACATGATCAGACCCATGATCAGAAATGGGATCATCTCCACCTGCTGGCTCTAAT 4189
 Db 4202 GCCACCATGATCCGGAAGGTGGACCAAAACGGAATCATCTCCACCTGCTGGCTCCAAT 4261
 Qy 4190 GATCTCACATCAGCCCGCCACTCAGCTGTGATTTCTGTCATGGATATTTCCAGGTAAGA 4249
 Db 4262 GACCTCACAGTGTCCGACCACTGAGCTGTGACTCGAGCATGAGCGTGGCCAGGTCCGT 4321
 Qy 4250 CTGAGTGGCCCAAGACTTAGCCATCAACCCAAATGGACAACTCACTTTATGTCCTGAC 4309
 Db 4322 CTAGAATGGCCGACAGACCTCGCCGTCAACCCCATGACAACTCCCTGTACGTTCTGGAG 4381
 Qy 4310 AACATGTGTCCTGCAAACTCTGAAAAACACAGGTGCGCATTTGCGCGGAGGCC 4369
 Db 4382 AACAACTCATCTTCGGATCAACGAGAACCAACAGGTGAGATCATCGCGGAGCGCCT 4441
 Qy 4370 ATGCACCTGCCAGTCCCTGCAATGACCACTTCTGCTAAGCAAGGTGGCCATCCACGCA 4429
 Db 4442 ATGCACCTGCCAGTTCCTGGGATCGACTACTTCCGCTCAGCAAACTCGGCATCCACTCT 4498
 Qy 4430 ACCCTGAGTTCAGCAGCGCTTGGCTGTTTACAAATGGGCTCTGTATATGTTGAG 4489
 Db 4499 GCGTGTGAATCAGCAGCGCATTTGCCATTTCTACACTGGGGTGTCTACATCACTGAG 4558
 Qy 4490 ACTGATGAGAAAAAGATCAACCGCATCAGGAGGTCAACCACTAGTGAGAGATCTCACTC 4549

Db 4559 ACGGACGAGAAAGATCAACGGCTAGCCCAAGTCAACCAATGGAGAGATCTGCCTC 4618
 Qy 4550 GTTGTGGGGCCCCAGTGGCTGTGATGTAAAAATGATGCCAACTGTGATTTGTTTCT 4609
 Db 4619 TTAGCCGGGGCGCTCAGACTGTGACTGCAAAAACGATGTCAACTGATCTGTACTCG 4678
 Qy 4610 GGAGACGATGTTATGCCAAGATGCAAGTTAAATACCCCATCTTCTTGGCTGTGTGT 4669
 Db 4679 GGAGATGACGTTAGCCACGAGCCCATCTGAACTGCCGCTCTCTTAGCCGTGGCT 4738
 Qy 4670 GCTGATGGGAGCTCTAGTGGCGACCTTGGGAACTCCGAAATTCGGTTTATCCGGAAG 4729
 Db 4739 CCGGATGGCAACCATCTAATTGACAGCTTGGGAATATCCGATCAGGGCGTCAACAA 4798
 Qy 4730 AACAGCCTTTCTCAACACCCAGAAATGATGAGCTGTCTTACCAATTTGACCAAGAG 4789
 Db 4799 AATAAACCCGTTCTTAACGCATTCACCAAGTATGAGGCTGATCTCCGGGAGAACAGAA 4858
 Qy 4790 CTCTATCTGTTGATACCAACCGGACCACTGTACACCCAAAGCTCCCCACAGGAGAC 4849
 Db 4859 TTGTACGTTTCAACGCTGATGTTATCCATCAGTACATGTGAGTCTGGTACTGGGAG 4918
 Qy 4850 TACTGTAACTTCACTCACTGAGGAGCGGAGATCACTCATCACAGACAACT 4909
 Db 4919 TACTGTAACTTCACTCACTGAGGCTGACAAATGACGTCAACGAGTTGATTTGACAAAC 4978
 Qy 4910 GGCACATGTTAAATGTCGCGGAGCTCTACTGGGATGCCCTCTGGCTGTGTGCTCCA 4969
 Db 4979 GGAATTTCCCTAAAGATCCGCGGAGAGAGTGGCATGCCCGCCACTGCTATGCGG 5038
 Qy 4970 GATGCCAGGTGTACTGGGTGACCATGGGACCAACAGTGCATCAAGAGTGTGACCA 5029
 Db 5039 GATAATCAGATTTATCACCTTACTTGTGGGACCAATGGAGGCTCAAAGCGGTGTCCACT 5098
 Qy 5030 CAGAGACAGAGTGGCCATGATGACATACATGGCAATTCGGCTTCTGGACCAAA 5089
 Db 5099 CAGAACCTGGAGCTGGGCTCATGACTTATGATGGGAACATGGACTCTTAGCCACCAAG 5158
 Qy 5090 AGCAATGAACCGGATGGACAACTTTTATGATACGACAGCTTGGCGCTGTGCAAAAT 5149
 Db 5159 AGTGATGAACCGGATGGACAACTTTTATGATACGACAGGCGCTGTGACCAAT 5218
 Qy 5150 GTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAAGTATACAGACAGTTCACTGGAT 5209
 Db 5219 GTGACCCGCCACCGGGGTGTGACCACTGTGACCGGGAATGGAGAAATCTATCACC 5278
 Qy 5210 GTCCAGGTAGAGACCTCCAGCAAGGATGAT--GTCCACATAACCCAACTGTCTGCC 5266
 Db 5279 ATTGATTTGAACTCCAAACCGGATGATGACGTCACTGTGATCACCACCTCTCTCC 5338
 Qy 5267 TCAGGCGCTTCTACACACTGTCAGACCAAGTCCGGAACAGCTACTATACATCGGGGCC 5326
 Db 5339 GTGGAGGCTCTCTATACAGTGTACAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 5398
 Qy 5327 GATGCTCTTGGCGCTGCTGTCGCAACCGCATGGAGTGGCGCTGCAGACTGAGCCC 5386
 Db 5399 AATGAAACCTTGGCGGTGATGTACGCCAACCGCATGGCTGTGAGTTTCAAGTGGGCC 5458
 Qy 5387 CACTTGTGGCTGGCACCGCTCAACCCCGCTGGGCAAGAGAAATGTCACTGTGCCATC 5446
 Db 5459 CAGCTCTCCGAGGACCACTCACCCCGCTCGGCGCTGCAACATCTCTCTGCCCATG 5518
 Qy 5447 GACACCGGCTCAACTGGTGGAGTGGCGCCAGCAAGAGAGAGTCTGGGGCCAGGTC 5506
 Db 5519 GAGAAATGGCTGAATCTCATCGAGTGGCGCTGAGGAGGAAACAGATCAAGGCAAGTCT 5578
 Qy 5507 ACTGCTTTGGCGCGCTGGGTGCAACCCGAAATCTCTATCTCTGACTTTGAT 5566
 Db 5579 ACATCTTTGGAGAGAGCTTCGGGTCCAGGAGGATCTCTCTGATGATTTATGAC 5638
 Qy 5567 CCGGTAAACCGCACAGAGAGATCTATGATGACCAACCGCAAGTTCACCTTGGATCTG 5626
 Db 5639 CGAAATATCCGTACGAGAGAGATCTACGATGACCAACCGCAAGTTCACCTTGGATCTC 5698

QY 5627 TACGACAGGCGGCGCGCCAGCCTCTGTGTCACCCAGCAGCAGGCTGAATGTGTCAAC 5686
DB |||||
QY 5699 TATGACAGGTGGCGCGCCCTTCTGTGGCTCCCGAGCAGTGGCTGGCAGCGGTCAAT 5758
DB |||||
QY 5687 GTGACATATCTCCCTCGGGGHTTCAITTCCTGGGATCCAGAGGGGCATCATGTCTGAAGA 5746
DB |||||
QY 5759 GTCTCTACTTCTTCAATGGCGGCTTGGCGGCTCCAGCGAGGGGCCATAGCAGAGG 5818
DB |||||
QY 5747 ATGAATACGACAGGCGCGCATCACATCCAGGATCTTCTGTATGGGAAGACATG 5806
DB |||||
QY 5819 ACAGACATTGACAGCAGCGCGATCTGTCCGCGATGTTCCCGACGGAAGTCTGG 5878
DB |||||
QY 5807 AGCTACACATATCTTAGAGAGTCCATGGTCTCTACTACAGCCAGAGGCGAGTATATC 5866
DB |||||
QY 5879 AGTTATCTTATCTTGACAAAGTCCATGCTCTTCTGCTACAGAGCAACGTCAGTACATA 5938
DB |||||
QY 5867 TTTGAGTTCGACAAAGATGACCGCTCTCTCTGTGACGATGCCAACGTGGCGCGCAG 5926
DB |||||
QY 5939 TTTGAATATGACTCTCCGATCGCTCCAGCAGTCACTATGCCAGTGTGCGCCGCGCAC 5998
DB |||||
QY 5927 ACATAGAGACCATCCGCTCAGTGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGC 5986
DB |||||
QY 5999 AGCATGTCCAGCACACTCCATTTGTTTACATCCGAAACATTTACAACCCACCCGAAGC 6058
DB |||||
QY 5987 AATGCTCAGTCATACAGGACTTCACTGAGGATGGGCACCTCTTCAACCTTCTACCTG 6046
DB |||||
QY 6059 AATGCATCGGTCACTTTGACTACAGTATGACGGCGGCATCTTAAGACATCTTTCTTG 6118
DB |||||
QY 6047 GGCACTGGCGCGAGGTGATATACAGTATGGCAACTGTCAAGCTGGCAGAGCGCTC 6106
DB |||||
QY 6119 GGCATGGCGCGAGGTGTTCTACAGTATGAAACTCTCCAGTTATACAGATAGTTC 6178
DB |||||
QY 6107 TATGACACCAAGTCAAGTTCACCTATGACAGAGCGCAGGCAATGTCTGAAGACCATC 6166
DB |||||
QY 6179 TACGACAGCAGCGGTGACCTTTGGTATGACGAGACCAACCGGTGCTCTGAAGATGTC 6238
DB |||||
QY 6167 AACCTACAGATGAGGCTTCACTCGACCATCGCTACCGTCAGATGGCGCCCTGATT 6226
DB |||||
QY 6239 AATCTCAAAGTGGGCTTCTCTGTACCATCAGGTACCGAAGGTTGGGCCCCCTTGTG 6298
DB |||||
QY 6227 GACGACAGATCTTCCGCTTCACTGAGGAAGGATGTCACACCCCGCTTTTGACTACAC 6286
DB |||||
QY 6299 GACAAGCAGATTTACAGGTTCTCTGAGGAAGGAATGATCAAGCCAGGTTTGATTAAC 6358
DB |||||
QY 6287 TAT---GACAACAGCTTCGGGTGACGAGTCCAGCATCGAGCTGTGATCAACGAGACCCACTG 6343
DB |||||
QY 6359 TATCAGACAAATAGCTTCCGCAATGCGCAGCATCAAAACCGGTCAATAGCAGACTCCCTTT 6418
DB |||||
QY 6344 CCATTTGATCTCTATCCCTATGATGTGTGAGGAAGACAGAGCAGTTTGGGAAGTTT 6403
DB |||||
QY 6419 CTTGTTGACCTTACCGCTATGACGAGATTTCCGGCAAGGTGAAACACTTCGGCAGTTT 6478
DB |||||
QY 6404 GGTGTATTTACTATGACATTAACAGATCATCACACAGCTGTCTATGACCCACCAACAG 6463
DB |||||
QY 6479 GGGGTCTACTACTACGACATCAACACAGATCATCACCAGTCCGCTCATGACGTTTAGCAAG 6538
DB |||||
QY 6464 CATTTTGATCATATGCGAGATGAAGAGTGCAGTATGAGATCTTCGCTCGCTCATG 6523
DB |||||
QY 6539 CACTTTGACACCCATGGGCGCATCAAGGAAGTGCATATGAGATGTTCCGTTCCCTCATG 6598
DB |||||
QY 6524 TACTGGATGACCGTCCAGTATGATAACATGGGCGGAGTGTGAGGAAGAGGCTGAAGTGA 6583
DB |||||
QY 6599 TACTGGATGACTGTGCAATATGACAGTATGGGTAGGTCATCAAGAGGGAACTGAAACTA 6658
DB |||||
QY 6584 GGACCTTACGCCAATACCACTCGCTACTCTATGATGATGATGCTGACGGCGAGCTGCAG 6643
DB |||||
QY 6659 GGGCCCTATGCCAACCAACAAAGTACACCTATGACTATGACGGGACCGGCGAGCTCCAG 6718
DB |||||
QY 6644 ACAGTCTCCATCAATGACAAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTG 6703
DB |||||
QY 6719 AGTGTGGCCGTCAATGACCGGCTACTCTGGCGCTATAGCTATGACCTCAATGGGAACCTG 6778
DB |||||

QY 6704 CACTTACTAGCCCTTGGGAACAGTGCAGCGGTCAACCACTTACGTTATGACATCCGGAC 6763
DB |||||
QY 6779 CACCTTCTAAACCCAGGAACAGTGTCTGCTCATGCTTACGCTATGACCTCCGTGAC 6838
DB |||||
QY 6764 CGCATCATCGCTCGGTGGTGAAGTGCATCAAGATGATGAGGATGCTTCTTCTGAGGAC 6823
DB |||||
QY 6839 CGGATAACCGGCTAGGGGACGTGACGTACAAAATCGATGACGATGCTATTTGTGTCAG 6898
DB |||||
QY 6824 CGGGCGGTGATATCTTTGATGACAACTCAGCTGGCTGTCTATCAAGGCTTCAAAACGG 6883
DB |||||
QY 6899 AGGGGTGACAGATCTTTTGAATACAACTCCAAAGGCTTCTGACGAGAGCATACAAAG 6958
DB |||||
QY 6884 GCTGGAGCTGGAGTGTGAGTACCGGTACGATGGCTGGGGCGGCGGTGTCAGCAAG 6943
DB |||||
QY 6959 GCCAGCGATGGAGCGTGCAGTACCGTATGACGAGTGGGCGCGGCTTCTTACAAG 7018
DB |||||
QY 6944 AGCAGCCACAGCCACCTCGAGTCTTCTATGACAGCTGACCAACCCACCAAGGTC 7003
DB |||||
QY 7019 ACCAACCTTGGGCCACCACTTACAGTACTTCTATCCGACCTCCACAACCCACAGTATC 7078
DB |||||
QY 7004 ACCCACTGTACAAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGCAGGA 7063
DB |||||
QY 7079 ACCCATGTTTAAACCACTCCCACTCTGAGATCACTCGCTCTACTATGACCTCCAGGC 7138
DB |||||
QY 7064 CACCTCTTGGCATGGAGCTGAGCAGTGTGATGATCAAGCAAGTTTACATAGCTTGTGACAAATC 7123
DB |||||
QY 7139 CACCTATTTGGCATGGAGAGCAGTAGTGGTGAAGATACTATGTGCTCTCAGAACACAG 7198
DB |||||
QY 7124 GGGACCCCTCTTGTGCTTTTAGTGAACAGGTTTGTATGATCAAGCAAAATCTGTGTACACA 7183
DB |||||
QY 7199 GGGACCCCTCTGCTGTGTACAGTATCAATGGCTCATGATCAAGCAAACTGCAGTACACA 7258
DB |||||
QY 7184 GCCTATGGGAGATCTACATGGATACCAACCCCACTTTCAGATCATATAGGTACCAT 7243
DB |||||
QY 7259 GCCTATGGGAGATCTACTATGACTCCAATCCAGACTTCCAGATGGTCAATGGCTTCCAC 7318
DB |||||
QY 7244 GGTGGCTTATGATCCACTCACCAAGCTTGTCCACATGGCGCGGAGATTTATGATG 7303
DB |||||
QY 7319 GGAGCCCTTATGACCCCTTACCAGCTCTGTCACCTTACTCACTGATTTATGAGCTG 7378
DB |||||
QY 7304 CTGGCCGAGACCTGGAATAGCCAGACCAAGCTGTGGAAGCACTTATGATGAGCAAC 7363
DB |||||
QY 7379 CTGGCAGACGCTGGAAGTCCCGCACTACACCACTGTGGAGGAACGT---GGGCAAGGAG 7435
DB |||||
QY 7364 GTCATGCTTTTATCTCTATGTTTCAAAAACCAACCCCATCAGCAACTCCAGGAC 7423
DB |||||
QY 7436 CGAGCCCTTCAACCTGTGATGTTCAAGAACAACTCTCTGAGCAATGAGCTGAC 7495
DB |||||
QY 7424 ATCAAGTGTCTTATGACAGATGTTTAAAGCTGGCTGCTCACTTTGATTTCCAGCTACAC 7483
DB |||||
QY 7496 TTAAAGAACTACGTGACAGAGCTGAAGAGCTGGCTTGTGATGTTGGATTTTCACTCAGC 7555
DB |||||
QY 7484 AACGTGATCTGTGTTATCCCAACACAGATGATGCCATGGAACCTCTCTTACGAGCTC 7543
DB |||||
QY 7556 AACATCATCTCTGATTTCCGAGAGCCAAATGTTTGTGCTCTCCCTCTATGA---- 7611
DB |||||
QY 7544 ATCCACACACAGATGAAAACGAGAGTGGGACAAACAGCAAGTCTATCTCTGGGGTACAG 7603
DB |||||
QY 7612 -----ACTGTGAGAGATCAAGCAAGCGAGACCGACAGCTCATTCAGGTGTCCAG 7663
DB |||||
QY 7604 TGTGAAGTACAGAAGCAGCTCAAGGCTTGTTCACCTTAGAACGGTTTGAAGCTCTAT 7663
DB |||||
QY 7664 CAGACAACTGAGAGGCATAACCGGCTTCTGCTCTGGAAGACAGGTCATCACTAAA 7723
DB |||||
QY 7664 GGCTCCCAATCACAGCTGCGAGAGGCTCCAAAGACCAAGAGTTTGTATCCAGCGGC 7723
DB |||||
QY 7724 AAGCTCATGCCAGCATCCGAGAAAGC-----AGGCCACTGGTTTGTATCCACACA 7777
DB |||||
QY 7724 TCAGTCTTTGGCAAGGGGTCAAGTGTTCCTTGAAGGATGGCGAGGTGACCAAGATC 7783
DB |||||
QY 7778 CCCATCATCGCAAGGCATCATGTTTGCCTATCAAGAGGGGGGTGACCAAGAGGTG 7837
DB |||||
QY 7784 ATCAGTGTGGCCATGAGGATGGCGAGGGTGTCTGCCATCTTGAACCATGCCACTAC 7843
DB |||||

[illegible]

Db	2198	GGTCTCAACTGTGAGCTGGCGAGGGTTCGAGTCCCGACACCAAGTGTATGTGGCGATGGCACT	2295
Qy	2150	TTCTCTCCCGGACACCGGGCTTTGACAGTGTGACCCAAAGCTGGACTGGACACGACTGTCTCT	2209
Db	2258	TACCTCCCTGACTCGGSCCTCTGCAGCTGTGATCCGAACCTGGATGGGTCCCGACTGCTCT	2317
Qy	2210	ATCGAGATCTGTGCTGCCGACTGTGTGGCCATGCGCTGTGCGTAGGGGGACACTGCCGC	2269
Db	2318	GTTG---TGTCTCAGTAGACTGTGGCACTCACGGCGTCTGCATCGGGGGAGCCTGCCGC	2374
Qy	2270	TGCGAGGATGGCTGGATGGGGGACCTCGACACAGCGGGCTGCCACCCGCGTGTGCC	2329
Db	2375	TGTGAAGAGGGCTGGACAGGCGCACTCTGTACCAGCGGTGTGCCACCCCGCTGCATT	2434
Qy	2330	GAGCANTGGAACTCCGCGGACGGCAAGTCGAGTGCAGCCCTGGCTGGAATGCGGAACAC	2389
Db	2435	GAGCACGGGACCTCTAAAGATGGCAAAATGTCAATGCCGAGAGGGCTGGAATGTGTGAACAC	2494
Qy	2390	TGCAACCATCGCTCACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGTTGTGCAAT	2449
Db	2495	TGCACCATTGAT-----GGCTGCCCTGATTTTGTGCAC	2527
Qy	2450	GGCAACGGCAGATGACTTTAGACCTGAAATGGTTGGCACTCGCTCTGCCAGCTGGCGTGG	2509
Db	2528	GGTAACGGGAGATGCACACTGGGTGAGAACAGCTGGCAGTGTGTCTGCCAGACCGGCTGG	2587
Qy	2510	AGAGGAGCTGGCTGTGACACTTCATCGGAGACTGCTCGGTGCAGACAGCAAGACAAATGAT	2569
Db	2588	AGAGGGCTGTGATGCAACAGTTGCCATGGAAACCTCTCGCCTGATAAACAAGGAATAATGAG	2647
Qy	2570	GGAGATGGCTGTGGGACTGCATGGACCCCTGACTGCTGCCCTCCAGCCCTGTGCCATATC	2629
Db	2648	GGAGATGGCTGTGGGACTGCTGGACCCCTGACTGCTGCCCTACAGTCAGCCTGTCAAGAC	2707
Qy	2630	AACCGCTGTGCTTTGGCTCCCTTAACCTCTCTGGAATCATCCAGGAGACACAGGTCCCT	2689
Db	2708	AGCCTGTCTCTCGCGGGGTCTCGGGACCCCTTGAGACATCATTCAGCAAGTCTTGGCAGCAAGGAC	2767
Qy	2690	GTGTCAACGACAGAACCTACACTCCTCTTATGACCGCATCAAGTTCTCTGTTGGCGCAGGAC	2749
Db	2768	TGGCCTGCAGTGA-----AGTCTCTATGACCGCATCAAGTCTTGGCAGCAAGGAC	2821
Qy	2750	AGCAGCACATAATCCCGGGGAGAACCCCTTTGTATGAGGGCATGCTTGTGTATTTCGT	2809
Db	2822	AGCACCCACATCATTCCTGGAGACAAACCCCTTCAATAGCAGCCTGTGTCTCTGATCCGA	2881
Qy	2810	GGCCAAGTGATGACATCAGATGGAAACCCCTCGTTGGTGTGTAACATCAGTTTGTGCAAT	2869
Db	2882	GGCCAAGTGATGACCTATGGATGGGACTCCCTTGGTGGGTGTGAATGTCTTTTGTCAAG	2941
Qy	2870	AACCTCTCTTTGGATATACAATCAGCAGGACAGATGGCAGCTTTGACTTTGGTGTGCAAAAT	2929
Db	2942	TACCCAAAATATGGCTACACCATCACTCGCCAGGATGGCACGTTTGACCTGATTGCCAAT	3001
Qy	2930	GGCGGATCTCCATCATCTCGGCTTCGAGCGGCGACCTTTTATATCACAGGAGGACACC	2989
Db	3002	GGGGGTTCTCCCTTGACTCTCTCACTTTGACGAGACCCCTTTTATGAGCAGGAGCGCAC	3061
Qy	2990	CTGTGGCTGCCATGGGATCGCTTCTTGTATGAGAAACCATCATCATGACATGAGGAG	3049
Db	3062	GTGTGGCTGCCATGGAAACACTTCTATGCCATGGAACACCTTGTATGAAGACCGAGGAA	3121
Qy	3050	AATGAGATTCCGAGCTGTGACCTGAGCAATTTTGGCGGCCCCAACCCGAGTGTCTCTCCA	3109
Db	3122	AACTCCATCCCCAGCTGTGACCTCAGTGGCTTTGTTCGGCCAGATCCAATCATCATCTCC	3181
Qy	3110	TCCCCACTGACGCTCTTCGCCAGCTCTGTGACAGAAAGGCCCCCAATTTGCCCGGAAATT	3169
Db	3182	TCTCTCTGTCCACTTCTTCAGGGCTTCCCCTGCTCGAAACCCCAATTGTGCTGAGACC	3241
Qy	3170	CAGCTTTGAGGAGGAATCTCTATCTCTGTGGCTGCAAGATGAGGTGAGTCACTTGAGC	3229
Db	3242	CAGGTTCTTCATGAAGAAATTGAGTCCCTGGTATCCCAATGTGAAGTCCGTTATCTCAGC	3301

QY 3230 AGCCGACCCCTGGCTACAAATCTGTCTCAGGATCAGCTCACCACCGGACCATCCCC 3289
 Db TCTAGAACTGCAGGGTATAAGTCGCTGCTGAAGATCACCATGACGAGTCCACAGTGC 3361
 QY 3290 TTCAACCTCATGAAGTGCACCTCATGTAGCGGTGAGGGCGCCCTCTTTCAAGAAAGTG 3349
 Db TTGAACCTCATCAGGGTTCACCTTGATGTTGCTGAGAGGGGCATCTCTCCAGAAAGTCA 3421
 QY 3350 TTGCTGCAGCCCGACGACCTGTCTATATTTATTTGGGACAAAGACAGAGCTCTACAAC 3409
 Db TTCCAGGCTTTCTCCCAACCTTAGCTACACATTCATCTGGGACAAAGACAGATGTTATGGC 3481
 QY 3410 CAGAAGGTGTTGGGCTTTTCAGAAAGCTTTGTTTCGCTGGGTTATGAATATGAATCCTGC 3469
 Db CAAAGGGTTATGGCCTATCGATGCTGTTGCTGCTGGTTGGAATATGAGACCTGC 3541
 QY 3470 CCGATCTAATCTGTGGGAAAAAGAAACAAAGTGTCTGAGGGCTATGAAATGAGCGG 3529
 Db CCCAGTCTCATCTGTGGGAAAAAGGACAGCCCTGCTTCAGGGATTCGAGCTGGACCTC 3601
 QY 3530 TCCAAGCTTCGAGGATGAGCCTAGACAAACATCATGCTCCCTCAACATTCAAAGTGTATC 3589
 Db TCCAACTTGGAGCTGCTGCTGGACAAACACACACCTCAATGTGAAACGGAAATA 3661
 QY 3590 CTGCACAAAGGGAATGGGGAAACCAAGTTTGTCTCAGCAGCTCTCCTGTCAATTGGGAGC 3649
 Db CTACACAAAGGACAGGGGAAACCAAGTTCTCTGACCCAGCAGCTGCGCATCAACGAGC 3721
 QY 3650 ATCATGGGCAATGGGCCCGGAGAGCATCTCTGCTGCCCGCAGCTGCAACGCCCTTGTGAC 3709
 Db ATCATGGGCAACGGTCCCGCAGAGCATCTCTGCTGCCCGCAGCTGCAACGCCCTTGTGAA 3781
 QY 3710 GGCACAAAGCTCTGGGCCCGCAGTGGCCCTCACCTGTGGCTCTACGGGAGCCTCTATGTG 3769
 Db GGCACAACTGTGTAGCCCTGTGGCCCTGGCTGTGGGATCGATGGGACCTTTGTT 3841
 QY 3770 GGTGATTTCACTACATAGAGATCTTCCCTCTGGAAATGTCAACAACTCCTAGAG 3829
 Db GGTGATTTCACTATATCCGGCGCATCTTCCCTCTCGAAATGTGACCAATCTTGGGAG 3901
 QY 3830 CTGAGGAATAAGATTTCAAGCATAGTCAAGTCCAGCACACAAATACCTACCTGGCCACA 3889
 Db TTACGAATTAAGATTTAAACATAGCAACAGCCAGACACAAAGTACTTCTGGCTGTG 3961
 QY 3890 GACCCCATGATGGGGCCGCTTCTCTTTCTGACAGCAACAGCCGGGGCTTTTAAATC 3949
 Db GACCCCATGATGGGCTCACTCTACGCTCTGACACCAACAGTCCGCAATCTACCGAGTC 4021
 QY 3950 AAGTCCACTGTGGTGAAGGACCTTGTCAAGACTCTGAGTGGTGGGGGACAGGT 4009
 Db AAGTCTCTGAGCGGAGCCAAAGACCTGGCTGGGAAATTCGGAATTTGGCAGGGACTGGC 4081
 QY 4010 GACCAGTGCCTCCCTTTGATGACACTCGCTGGGGGATGGTGGGAGGCCACAGAGCC 4069
 Db GAACAATGTCTACCTTTGATGAGCCGCTGTGGGATGAGGGAAGGCTGTGGAGCC 4141
 QY 4070 ACATCAACAAATCCAGGGTATTACAGTGGACAAAGTTTGGGCTGATCTACTTTCTGGAT 4129
 Db ACCCTGATGAGCCCGCAGAGGTATTGCAAGTGAACAAGAAATGGGCTTATGTACTTTGTTGAT 4201
 QY 4130 GGCACCATGATCAGACGATCGATCAGATGGGATCATCTCCACCTGCTCGGCTCTAT 4189
 Db GCCACCATGATCCGGAAGGTGGACCAACGGAAATCATCTCCACCTGCTGGGCTCCAAAT 4261
 QY 4190 GATCTCATCAGCCCGCCACTCAGCTGTGATTTCTGATGATATTTCCAGGTAAGA 4249
 Db GACCTCAGCTGTCCGACCACTCAGCTGTGACTCGAGCATGGACGTGGGCCAGGTCCTG 4321
 QY 4250 CTGAGTGGCCCAAGACTTAGCCATCAACCCATGAGACACTCACTTTATGTTCTCGAC 4309
 Db CTAGAATGGCCGACAGACCTCGCGCTCAACCCCATGAGCAACTCCCTGTACGTTCTGGAG 4381

QY 4310 AACAAATGTGGTCTCTGAAAACCAACAGGTGCGCATTTGTCGCGGGAGGCC 4369
 Db AACAAATGTCTCTGCGGATCAGGAGAAACCAACAGGTGAGCATCATCGGGAGGGCT 4441
 QY 4370 ATGCACTCCAGGTCCTGGCAATTCACATTCCTGTTAAGCAAGGTGGCAATCCACGCA 4429
 Db ATGCACTCCAGGTTCCCGGATCGACTACT---CGCTCAGCAAACTCGCCATCCACTCT 4498
 QY 4430 ACCCTGGAGTCAGCACCGCTTTGGCTGTTTTCACAAATGGGCTCTATATTCCTGAG 4489
 Db GGGCTGGAAATCAGCAGCGCCATTCCTCAITTTCTCAGCTGGGCTCTCATCATCTGAG 4558
 QY 4490 ACTGATGAGAAAAAGATCAACCGCATCAGGAGGTCAACCTAGTGGAGAGATCTCACTC 4549
 Db ACGGACGAGAGAAGATCAACCGCTACGCAAGTCAACCAATGGAGAGATCTGCTC 4618
 QY 4550 GTTGTGGGGCCCCAGTGGCTGTGACTGTAAAATGATGCAACTGTGATGTTTCT 4609
 Db TTAGCCGGGGCGGCTCAGACTGTGACTGCAAAAACGATGTCAACTGTCTACTCTCG 4678
 QY 4610 GAGACGATGTTATGCCAAGATGCAAAAGTTAAATACCCCATCTTCTTGGCTGTGTGT 4669
 Db GGGATGAGCTTTAGCCACCGGACGCTCTGAACTCGCGCTCTCTCTTAGCGTGGCT 4738
 QY 4670 GGTGATGGGAGCTCTAGTGGCGGACCTTGGGAAACATCCGAAATTCGGTCTTATCCGGAAG 4729
 Db CCGGATGGCAACCATCTACATTCGACACCTTGGGAATATCCGATCAGGGCGGTGAGCAAA 4798
 QY 4730 AACAGCCTTTCTCTCAACCGCAGCAATGATGAGCTGTCTTCCAAATGACCAAGAG 4789
 Db AATAAACCCGTTCTTAAAGCATTAACAGATGAGGCTGATCTCCGGGAGAACAGAA 4858
 QY 4790 CTCTATCTGTTTATACCAACCGGCAAGCCTGTACACCAAGCCTGCCACAGAGAC 4849
 Db TTGTACGTTTCAACGCTGATGGTATCCATCAGTACACTGTGAGTCTGGTGTGGGAG 4918
 QY 4850 TACCTGTACAACTTCACTTACACTGGGACGGGACATCACTCATCAGACAACTAAT 4909
 Db TACTTGTACAAATTCACATACAGCCTGACAAATGACGTCAGCGAGTTGATTGACAACT 4978
 QY 4910 GGCACATGTTAAATGTCCGCGAGACTCTACTGGGATGCCCTCTGGTGGTGTCCCA 4969
 Db GGGATTTCCCTAAAGATCCGCGGACAGCAGTGGCATGCCCGCCACCTGTCTATCGCG 5038
 QY 4970 GATGCCAGGTGATCTGGGTGACCATGAGGACCAACAGTGCATCAAGAGTGTGACACA 5029
 Db GATAATCAGATTTATCACCTTACTGTGGCCACCAATGGAGGCTCAAAAGCCTGTCTCACT 5098
 QY 5030 CAAGGACAGGTTGGCCATGATGATACATGCAATTCGCGCTTTCTGGCAACCAAA 5089
 Db CAGAACCTGGAGCTGGGCTCATGACTTATGATGGAAACATGGAATCTTAGCCACCAAG 5158
 QY 5090 AGCAATGAAAAAGGATGGACAACTTTTATGATGATGACAGCTTTGGCGCTGTGACAAAT 5149
 Db AGTATGAAAAAGGATGGACAACTTTTATGATGATGACAGGCGCTGTGACCAAT 5218
 QY 5150 GTGACCTTCCCTACTGGCCAGGTGAGCAGTTTCCGAGTGATACAGACAGTTCAGTGAT 5209
 Db GTGACCCCGCCACAGGGGTGGTGACCACTGTGACCGGAAATGGAGAAATCTATCACC 5278
 QY 5210 GTCCAGGTAGAGACCTCCAGCAAGGATGAT---GTCAACATAACCAACCACTGTCTGCC 5266
 Db ATTGACATGAGAACTCCAAACCGGATGATGAGCTCACTGTGATCAACCACTCTCTCC 5338
 QY 5267 TCAGCGGCTTTTACACTGTGTGCAAGCAAGTTCGGAAACAGTCTACTATACATCGGGGCC 5326
 Db GTGAGGGCTCTTATACAGTGGTCAAGATCAAGTGCAGAAACAGCTACAGCTCTGCAAT 5398
 QY 5327 GATGCTCTTCTGGCTGTGTCGCCCAAGGATGGAGGTGGCGCTGCAGACTGAGGCC 5386
 Db AATGAAACCTCGGGGTGATGTACGCCCAAGCGCATGGCTGTGAGCTTCCACAGTGGGCC 5458
 QY 5387 CACTTCTGCTGGCACCCGTCAACCCCAAGCGGTGGGCAAGAGGATGTCAAGTGTGCCCATC 5446

Db 5459 CACGTCCTCGCAGGCACCATCACCCACCATCGGGCGCTGCAACATCTCTCTGCCCCTG 5518
Qy 5447 GACAAACCGCTCAACCTGTTGGAGTGGCGCCAGGCGCAAGAGCAGGCTCGGGCCAGGTC 5506
Db 5519 GAGAAATGCGCTGAACCTCACTGAGTGGCGCTGAGGAGGAAACAGATCAAGAGGCAAGTC 5578
Qy 5507 ACTGCTTTGGCGCCCGCTGCGGGTGCACAAACGAAATCTCTATCTCTGGACTTTGAT 5566
Db 5579 ACCATCTTTGGAGGAGCTTCGGGTCCAAGGAGGATCTCTGTCATTTGATTATGAC 5638
Qy 5567 CGGTAAACAGCAGAGAGATCTATGATGACACCGCAAGTTACCCCTTCGGATCTTG 5626
Db 5639 CGAAATATCGGTACGAGAGAGATCTAGATGACACCGGAAATTCACCCCTGAGGATCATC 5698
Qy 5627 TACGACAGCGCGCGCGCCAGCCTCTGTGTCACCAAGCAGAGGCTGAATGTTGTCAC 5686
Db 5699 TATGACCAAGTGGCCCGCCCTCTCTGTGGCTCCGAGAGTGGCTGCGACCGCTCAAT 5758
Qy 5687 GTGACATACCTCCCTGGGGTTACATTTGTCATCCAGAGGGGTCATGTCGAAAGA 5746
Db 5759 GTCTCTTCTTCAATGGCGCTTGGCCGCTCCAGGAGGGCCATGAGCGAGAG 5818
Qy 5747 ATGGAATACGACAGCGCGCGCCGATCATCCAGGATCTTCCTGATGGGAGACATGG 5806
Db 5819 ACAGACATGACAAAGCGCGGATCGTGTCCGCAATGTCGCCAGCGGAAAGTCTGG 5878
Qy 5807 AGCTTACATACCTTTAGAGAGTCCATGGTCTGCTACTACAGCCAGAGGCGAGTATATC 5866
Db 5879 AGTTATTCCTATCTTGACAGTCCATGCTCTCTGCTACAGAGCCACGTCAGTACATA 5938
Qy 5867 TTTGAGTTCGACAAAGATGACCGCTCTCTCTGTGACGATGCGCCAAAGTGGCGCGAG 5926
Db 5939 TTTGAAATGACTCTCCGATCGCTCCAGCGAGTCACTATGCCCAGTGTGCGCGCGAC 5998
Qy 5927 ACCTAGAGACCATCCGCTCAGTGGGCTACTACAGAAACATCTACAGCCCTGAGGCG 5986
Db 5999 AGCATGTCCAGCACACTTCATGTTGTTACATCGAAACATTTACACCCACCGAAAGC 6058
Qy 5987 AATGCTCAGTACAGGACTTCACTAGGATGGGCACTCTTTCACACCTTCTACCTG 6046
Db 6059 AATGATCGGTCTCTTTGACTACAGTACAGCGCGCATCTTAAAGACATCTTCTCTG 6118
Qy 6047 GGCACTGGCGCAGGTTGATATACAGTATGGCAACTGCAAGCTGGCAGAGGCTC 6106
Db 6119 GGCACTGGCGCAGGTTGTTCTCAAGTATGGAAACCTTCCAAAGTTATCAGAGATAGTC 6178
Qy 6107 TATGACACCAAGTTCAGTTCATCCTATGAGAGAGGATGGTCAAGCCGTTTGTACTACAC 6166
Db 6179 TAGCAGAGCAGCGCTCACCTTTGGTATGACAGAGCCACCGTGTCTGAAGTGGTC 6238
Qy 6167 AACCTACAGATGAGGCTTCACTGACCATCGCTACCGTCAGATTTGGGCCCTGATT 6226
Db 6239 AATCTCAAAGTGGGCTTCTCTGTACCATCAGGTACGAAAGTGGGCCCTTGTG 6298
Qy 6227 GACCGAGATCTTCCGCTTCACTGAGGAGGATGGTCAAGCCGTTTGTACTACAC 6286
Db 6299 GACAGCAGATTTACAGTTCTCTGAGGAGGATGATCAACCGCAGTTTGTATATAC 6358
Qy 6287 TAT---GACAAAGCTTCCGGTGAACAGATCGAGGCTGTGATCAACAGAGCCCACTG 6343
Db 6359 TATCAGACAAATAGCTTCCGATTTGCGATGACGATCAAAACCGCTATTAGCGAGACTCCCTT 6418
Qy 6344 CCATTTGATCTATCGCTATGATGTTGTCAGGCAAGACAGAGCAGTTTGGAGGTTT 6403
Db 6419 CTTGTTGACCTTACCGCTATGACGAGATTTCCGCAAGGTGGAACATCTTCGCAAGTTT 6478
Qy 6404 GGTGTCTATTTACTATGACATTAACAGATCATCACCAGCTGTCTATGACCCACCAAG 6463
Db 6479 GGGGTCTATCTACTACGACATCAACAGATCATCACCACTGCCGTATGACGCTTAGCAAG 6538
Qy 6464 CATTTTATCATGTCAGGATGAGGAGTCCAGTATGAGATCTTCGCTCGCTCATG 6523

Db 6539 CACTTTGACACCCATCGGGCGCATCAAGGAAGTGCATATAGATGTTTCCGCTCCCTCATG 6598
Qy 6524 TACTGGATGACCGTCCAGTATGATAACATGGGCGAGTAGTGAAGAGGAGCTCAAGGTA 6583
Db 6599 TACTGGATGACTGTGCAATATGACAGTATGGGTAGGTCATCAAGAGGGAACCTGAAACTA 6658
Qy 6584 GGAACCTTACGCAATACACCTCGCTACTCTCTATGAGTATGTCGACGCGCCAGCTCCAG 6643
Db 6659 GGGCCCTATGCAACACACCAAGTACACCTATGACTATGACGGGAGCGCCAGCTCCAG 6718
Qy 6644 ACAGTCTCCATCAATGACAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTG 6703
Db 6719 AGTGTGGCGCTCAATGACCGGCTTACCTGGCGCTATAGCTATGACCTCAATGGGAACCTG 6778
Qy 6704 CACTTACTGAGCCTCGGAAACAGTGCAGGCTCACACACTACGCTATGACATCCGCGAC 6763
Db 6779 CACTTCTAAACCCAGGAAACAGTGTCTGCTCATGCCCTTACGCTATGACCTCCGTGAC 6838
Qy 6764 CGCATCACTCGGCTGGGTGACGTGCAATACAAGATGATGAGGATGCTTCTGAGCGAG 6823
Db 6839 CGGATAACAGGCTAGGCGAGCGTGCAGTACAAATCGATGACGATGGCTATTTGTGCCAG 6898
Qy 6824 CGGGCGGTGATATCTTTGAGTACAACCTCAGCTGCGCTGCTCATCAAGGCTTACAACCGG 6883
Db 6899 AGAGGTGAGACATCTTTGAAATCAACTCCAGGGGCTTCTGAGAGAGATACACAAG 6958
Qy 6884 GCTGGAGCTGGAGTGTGAGTACCGTACGATGCGCTGGGGCGCGCTGTCCAGCAAG 6943
Db 6959 GCCAGCGATGGAGCGTGCAGTACCGCTATGACGAGTGGCGCGCGGCTTCTTACAAG 7018
Qy 6944 AGCAGCCACAGCCACCTGCACTTCTTATGACAGCTGACCAACCCCAACCAAGGTC 7003
Db 7019 ACCAAGCTGGCGCCACCTACAGTACTTCTTACCGACCTCCACAACCCCAACAGTATC 7078
Qy 7004 ACCAAGCTGTACAACCACTCCAGCTCTGAGATCACTCCCTCTACTACGACTTGCAGGA 7063
Db 7079 ACCATGTTTACACCACTCCAACCTCTGAGATCACTCGCTCTACTATGACCTCCAGGC 7138
Qy 7064 CACTCTTTGGCCATGAGAGCTGAGCAGTGTGATGAGTTTTACATAGCTTTGTGCAACATC 7123
Db 7139 CACTATTTGGCCATGAGAGCAGTGTGAGTGTGAGATACTATGTGCTCAGACCAACG 7198
Qy 7124 GGGACCCCTCTGTGCTTTTGTAGTGAACAGGTTTGTATGATCAAGCAATCTCTGTACACA 7183
Db 7199 GGGACCCCTCTGTGCTGTGACAGTATCAATGGCTCATGATCAAGCAACTGAGTACACA 7258
Qy 7184 GCCTATGGGAGATCTACTATGACTCCAATCCAGACTTCCAGATGGTCAITGGCTCCAC 7243
Db 7259 GCCTATGGGAGATCTACTATGACTCCAATCCAGACTTCCAGATGGTCAITGGCTCCAC 7318
Qy 7244 GGTGGCTCTATGATCACTCAACAGCTTGTCCATGCGCGCGGAGATTATGATGTG 7303
Db 7319 GGAGGCTCTATGACCCCTTCCCAAGCTCGTCCACTTACTCAACGTGATTATGAGCTG 7378
Qy 7304 CTGCGCGAGCCTGAGCTAGCCAGACAGAGCTGTGGAAGCAGCTTACTAGCAGCAAC 7363
Db 7379 CTGCGAGAGCGTGGAGCTGCCCGACTACACCATGTGGAGAACGT---GGCAAGGAG 7435
Qy 7364 GTGATGCTTTTATCTCTATATGTTCAAACCAACCAACCCCATCAGCAACTCCAGGAC 7423
Db 7436 CCAGCCCTTCAACCTGTATCATGTTCAAGAACCAACCAATCTCTGAGCAATGAGCTGAC 7495
Qy 7424 ATCAAGTCTTCAAGCAGATGTTTAAAGCTGGCTGCTCACTTTGGATTTCAGCTACAC 7483
Db 7496 TTAAGAACTACGTGAAGAGCTGGCTGTGATGTTGATTTGATTTTCACTCAGC 7555
Qy 7484 AACGTGATCCCTGTTATCCAAACAGAGATGATGCCATGGAAACCTCTCTACGAGCTC 7543
Db 7556 AACATCATCTCTGATTTCCGAGAGCCAAATGTTATTTTGTGCTCTCCCTCTATGA--- 7611
Qy 7544 ATCCACACAGATGAAACCGAGAGTGGGACAAACAGCAAGTCTATCTCTCGGGGTACAG 7603
Db 7612 -----ACTGTCAAGAGTCAAGCAAGCGAGAACGACAGCTCATACAGGTGTCCAG 7663

QY 7604 TGTGAGTACAGAGCAGCTCAAGGCGCTTTGTCTACCTTAGAACGGTTTGACAGCTCTAT 7663
Db 7664 CAGACACTGAGAGCATACACAGGCTTCTGCTCTGGAAGACAGGTCATCACTAAA 7723
QY 7664 GGTCTCAACATPACACAGCTGCCAGAGCTCTCAAGACCAAGAAAGTTTGATCCAGCGGC 7723
Db 7724 AAGCTCCATGCGACATCCGAGAGAAAGC-----AGGCCACTGGTTTGTACCAACACA 7777
QY 7724 TCAGTCTTTGCAAGGGGTCAAGTTTGCCTTTGAAGGATGCCAGTGCACACAGATC 7783
Db 7778 CCCATCATCGCAAGGATCATGTTTGCCTCAAGAAAGGCGGGTGCACACAGAGAGT 7837
QY 7784 ATCAGTGTGGCAATGAGGATGGGCGAAGGGTCTGCTCCATCTTTGAACCATGCCACTAC 7843
Db 7838 TCTAGCATCGCAGTGAAGGACAGCGCAAGTASCATCCGTTGTGAACATGCTACTAC 7897
QY 7844 CTAGAGAACTGCACTTACCAATGATGGGGTGATACCATTTACTTTTGAACACAGGA 7903
Db 7898 TTAGACAGATGCACTACAGATCGAGGCAAGGACACACACTACTTTTGTGAAGATCGC 7957
QY 7904 CTTTCAAGAGTGACCTGGCCATCTGGGCTCAGTGGGGGGCGGAACCTTGGAGAAT 7963
Db 7958 GCCCGGATGGTACCTGGTCACTAGGAACACCATTTGGGCGCAAGGTGTGGAGAT 8017
QY 7964 GGGTCAACGTCACTGTGTCAGATCAACACAGTACTTAATGSCAGGACTAGACGCTAC 8023
Db 8018 GGGGTGAACGTGACGGTGTACAGCCCAACGCTGCTGCTGAATGGCAGGACTCGAAGGTT 8077
QY 8024 ACAGACATCCAGCTCCAGTACGGGCACTGTGCTTTGAACACAGCTACGG-----G 8074
Db 8078 ACCAATGATGTTCCAGTACTCCAGCTGTGCTCAGTATCCGCTACGGCTACACCC 8137
QY 8075 ACAAGTGTGATGAGGAGAGGACAGGCTCTGAGACAGGGGCTGCGGACAGAGCGCTCGC 8134
Db 8138 GACACGCTGAGAGAGAAAGGCGCGCTCTGGAACCAAGGGGACAGAGCGCTGGT 8197
QY 8135 CAAGCTGGGCGCGGACAGACAGTCTGAGACAGGGGCTGCAAGGCTACAGCGGTTT 8254
Db 8198 ACTCCCTGGGCAAGAGGACAGAAAGCCAGGACGCGGAGAGGCGCGCTGTGG 8257
QY 8195 ACAGAGGGGAGAGCAGAGCTCTGAGACAGGGGCTGCAAGGCTACAGCGGTTT 8254
Db 8258 ACGAGGGGAGAGAGCACTCTGAGACAGGGGCTGCAAGGTTATGAGGCTAT 8317
QY 8255 TTCTGTACTCTGTGAGAGTACCCAGAACTGTACAGAGCGGCAACACATCCACTTC 8314
Db 8318 TACGTACTTCCGGTGGAAACAGTACCCGAGCTGGCAGACAGTAGCAGCAACATCCAGTTC 8377
QY 8315 ATGAGACAGAGAGAGTGGCGGAGGTGACAGA 8348
Db 8378 TTAAGACAGATGAGTGGGAAAGAGGTAAACAA 8411

RESULT 15
US-09-800-198-62
; Sequence 62, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrmann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 8797
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-800-198-62

Query Match 40.5%; Score 3384.8; DB 10; Length 8797;
Best Local Similarity 64.8%; Pred. No. 0;
Matches 5467; Conservative 0; Mismatches 2712; Indels 255; Gaps 20;

QY 35 ATGGAGCTGAAGGAGAGAGGAGGCTTTACCGCTCGCTGACCCGCGCGCGGACCGAGCGC 94
Db 113 ATGGATGTAAGGACCGGGGACATCGCTCTTTGACAGGGGACGGTGTGGCAAGAGTGT 172
QY 95 CGCTACACAGCTCGTCCGCGACAGCGAGGAGGGAAG-----CCCGCAGAAATTCGTAC 151
Db 173 CGCTACACAGCTCGTCTCTGACAGTGAAGTCCGCTGTGCCCACTCAGAAGTCTTAC 232
QY 152 AGCTCCAGGAGACCTGTAAGGCTTACGACAGGAGCGCGGCTAGCCTATGCGACCGC 211
Db 233 AGTTCCAGTGAAGCTTTGAGGCTTATGACCATGACAGCAATGCACTATGGAACCGA 292
QY 212 GTCAAGGACATTTGTCGCGGAGGCGGAGGAATTTGCGGCACAGGTGCCAATTCAC 271
Db 293 GTCAAGGACATTTGTCGCGGAGGCGGAGGAATTTGCGGCACAGGTGCCAATTCAC 352
QY 272 CTGCGGAGCTCGGGCTGGAAGTAAAGCCCTCAGCGGACCTGTACCGGACAGAC 331
Db 353 CTGGCAGAAATTTGGGAATCTGCGAG-----CCCTCCACACCGAAGTGTGTCTCGAC 409
QY 332 ATTGGCTGCCCCAATGCGGCTACTCCATGCGGGGCTGCTGTGATGCCGACATGAGGCT 391
Db 410 ATGGGTATCTCCACAGGGCTACTCCCTGAGCACTGGGTCTGTATGCACTCGGACAC 469
QY 392 GACACGGTCTCTGCGCTGAGCACCGCGTCTGTGGGCGCGGAGACACAGGTCAGG 451
Db 470 GAGGAGGAGTGTCTCCAGAACATGCCATCAGACTGTGGGAGGAGGGAATAAATTCAGG 529
QY 452 CGCAGCTCTGCTGCTGCGAGCGGGCCAAATCCAAATCTCACACTCACCGACACCGAGCAT 511
Db 530 CGCAGCTCTGCTGCTGCGAGCGCGGAGAACTCGGCGCTTACTCTGACTGACTGACAA 589
QY 512 GAAACACTGAGACTGATCATCCGGGCGGCG-----TGCAAGAACCAACGCGGCTCGGAC 567
Db 590 GAAATTAATTCGATGACGACAAATGCTGCTCCCAATCCCACTACATCTCTGCTAGGCTC 649
QY 568 GCGCGCGCGCGCTCTGCGACCGCCACACCGCCCAACCGACACACCGCGGCT----- 620
Db 650 CTCCCATCTGCTCAGCTGCTAGCTCCCAATATCTCCACAGTGTAGTGCAGATGCCA 709
QY 621 ----- 620
Db 710 TTGCTAGACAGCAACACCTCCCATCAGATCATGAGACCAACCCCTGATGAGGAATTCCT 769
QY 621 -CCATTAATCTGTAACCGGGGCACTTACGCGGAGGAGCAACCGGACCGGCGGCGCC 679
Db 770 CCAATTTCACTCTCTCAGAGCATGCTCAGGGGCGGCGGAGGAGGCTCCAGAGTGGGCT 829
QY 680 ACGGACCGC-----TCGCTCTCCGAGAGCGGCGGCTGCGGCGGCGGCG 721
Db 830 CCAACACACACAGCGAGTCAACACTGAGGCGGCTCTGCCACCGGCTCATTAACACAC 889
QY 722 CAGGAGCTGCGGCGGCGGAGGAACTGGGTGTCTCAACAGCAACATCCCTCTGAGAC 781
Db 890 CTGTCCACACCACTCTCTCGGCGCAACTCCCTCAACAGAGAACTCACTGACCAATCGG 949
QY 782 AGGAACCTAGGCAAGCAGCATTCCTAGGACATTTGAGGACAACTC-----ATT 832
Db 950 AGTCAATCCAGCGCGGCGGCTCTGCGGCGGAGAGCTTGGCGGCGGCGGAGTGTGT 1009
QY 833 GAGATGACATTTCTCGGCGGCTCCCGGCGGCTTACGATGAGGCGGCACTTCCTC 892

Db	1010	CAGCTCCAGGATAGCTGGTGCTGAAACAGTAGTAACCTCCACCTGGAGACTCGCACTTCCTT	1069
Qy	893	TTCAA---GCCTGGAGGACCTCCCGCTCTTCTTGCAACCATCACAGGGTACCACACTG	949
Db	1070	TTCAAACGTCGTCGGAGACAGCCCTGTTACGAGCTCTTCTCGGGATACCCCTTG	1129
Qy	950	ACGTCACACAGTGTACTCTCTCCGCCCGACCCCTGCGCCCGCAGACACTTCGCGCGG	1009
Db	1130	ACCTCAGGGACCGTTTATACACCACACACCCCGCTGTGCGACCGGAATACATTTCTCCAGG	1189
Qy	1010	CCGCGCTTTAACTCAAGAGCCCTCCAAAGTACTGTAACTGGAAGTGGCAGACCCCTGAGC	1069
Db	1190	AAGGCTTTCAAGCTGAAGAAACCTCCAAATACTGCAAGTTGGAATGTGCTGCCCTGTCT	1249
Qy	1070	GCCATCGTCACTCAGCCACTCTGGTCACTCCTGCTGTCATACTTTGTGGCCATTTTGTG	1129
Db	1250	GCCATCGCGCGCCCTCCTCTTGCCCATTTTGTGTCATATTTATAGCAATGCACTCG	1309
Qy	1130	TTTGGCTAAACTGGCACCCTGCAGCCGATGGAGGGGAGATGATAGATACCGAGGAC	1189
Db	1310	CTCGGACTCAATTGGCACTCCAGCCCGCAGATGGACACACCTTTTAA-----ATGGC	1363
Qy	1190	ACGCCAGCAGTTGGCTGTGCCAACCGACGCTCTCCATACCCCTCAGGGGCACCTGGC	1249
Db	1364	GTAAGGACCGCTTACCAGGAAACGATGATGTGCAACAGTGCATCTGGAGGCAAGTG	1423
Qy	1250	TTAGAGACCCCTGACAGAAAGGCAAGAAACACAGAAAGGCCAGCCAGTAGTTCTTTT	1309
Db	1424	CCCTGGTCATTGAAAA-----	1439
Qy	1310	CCAGAGACAGTTTCATAGATTTCTGGAGAAATTCATGTGGAAGCGAGCCTCCAGAAG	1369
Db	1440	-----ACAGCAGCATAGACAGTGGCGAGCAGAGTTGTCGCGGGTGACACAGAA	1492
Qy	1370	ATTCCTCCTGGCACTTTCTGGAGATCTCAAGTGTTCATAGACCACTCTGTGCATCTGAAA	1429
Db	1493	GTCCACACAGGGGTGTTTGGAGTCCAGATTCCATTCAGTCAGCCTCAATCTTAAAG	1552
Qy	1430	TTCAATGTCTCTGGNAAGCAGCCCTGGTTGGCAATTTATGCGAGAAAGGCCTCCCT	1489
Db	1553	TTCAACATCTCCCTGGGCAAGATGCCCCCTCTTCGGTGTCTATATAGGAGAGGACTACCA	1612
Qy	1490	CCTTTCACATACACAGTTTGACTTTGTGGAGCTGTGGATGGCAGGAGGCTCTTAACCCAG	1549
Db	1613	CCGTCTATGCCAGTAGATCTTCATGGAAGCCCTGGATGGAA-----AG	1657
Qy	1550	GAGCGCGAGCCTTAGAGGGACCCCGCGCAGTCTCGGGAACTGTGCCCCCTCCAGC	1609
Db	1658	GAGAAATGGAGCGTGGTCGAGTCCCGCAGGGAACCGCGAGCATCCAGACTCTGGTGAG	1717
Qy	1610	CATCAGACAGGCTTCATCCAGTATTTGGATTTCAGGAATCTGGCACTTGGCTTTTACAAT	1669
Db	1718	AACGAGGCTGTGTTGTGCAGTACTTGGATGTGGGCTGTGGCACTTGGCCTTCTACAT	1777
Qy	1670	GACGGAAGGAGTCAGAAGTGGTTTCCTTTCTACCACTGCCATGCAATGAGTCGGTGGAATAC	1729
Db	1778	GACGGCAAGGACAAAGGAGATGGTCTCTTCAACACTGTGTCTTAGATTTCAGTGCAGGAC	1837
Qy	1730	TGCCCGAGCACTCTATGSCAATGTGTACTGATCATCTCGGACCTGCCACTGCTTCCTG	1789
Db	1838	TGTCACGGAACCTTCAGGGAACGGTGAATGCGGTGTCTGCAATGTCTACTGTTTCCCCA	1897
Qy	1790	GGTTTCTGGCCCCGACTGTGGCAGAGCCCTCTGSCCCCGTGTCTGTAGCGAAATGGC	1849
Db	1898	GGATTCTTAGTGCAGACTGTGCTAAAGCTGCCTGCTGTACTGTSCAGCGGAAATGGA	1957
Qy	1850	CAATACATGAAGCAGATGCTTGTGCCAGTGGCTGGAAAGCGCTGAGTCGGATGTG	1909
Db	1958	CAGTATTCTAAAGGAAAGTCCAGTGTACAGCGGCTGGAAAGGTGAGAGTGTGATGTG	2017
Qy	1910	CCCACCAACAGTGTATCGATGTGGCTGTGACGACCAACCATGCGCACTGCATCACGGGCAAC	1969

Db	2018	CCTATGAACCAATGTATCGATTCCTTCTGTGGGGCCCATGGCTCCTGCATTGATGGGAAC	2077
Qy	1970	TGATCTGCAACCCCTGGCTACAAGGGCGAGAGTGTGAGGAAGTGGACTGCATGGACCCC	2029
Db	2078	TGCGTGTGCTGCTGTGCTACAAGGGCGAGCACTGTGAGGAAGTTGATTGCTTTGGATCCT	2137
Qy	2030	ACATGTTACGGCCGGGTGTCTGGGTGAGAGGCGAATGCCATTGCTTTGTGGGATGGGA	2089
Db	2138	ACCTGCTCCAGCATTGGTGTCTGTGTGAATGSAAGTGTCTATGCAACCCCGGCTGGGGT	2197
Qy	2090	GGCACCAACTGGCAGAGACCCCGAGGCCACATGCTTAGACCAGTGTTCAGGCCACGGAACC	2149
Db	2198	GGTCTCAACTGTGAGCTGGCGAGGGTCCAGTGGCCAGACCAGTGTAGTGGCATGGCACT	2257
Qy	2150	TTCCCTCCCGACACCGGGCTTTTGAGGTGTGACCCAAAGCTGGACTGAGACAGACTGTTCCT	2209
Db	2258	TACCTCCCTGACTCCGGCCCTCTGCAGCTGTGATCCGAATCTGGATGGGTCCCGACTGCTCT	2317
Qy	2210	ATCGAGATCTGTGTCGCCGACTGTGTGGGCATATGCGCGTGTGCGTAGGGGACACTGCCGC	2269
Db	2318	GTTG- --TGTCTCAGTAGACTGTGGCACTCAGCGGCTGTGCATCGGGGAGAGCTGCCGC	2374
Qy	2270	TGCGAGGATGGCTGGATGGGGGGCAGCTCGCACAGCGGCCTGCCACCCGCGCTGTGCC	2329
Db	2375	TGTGAAGAGGGCTGGACAGGCGCAGCTTGTGACAGCGCGTGTGCCACCCCGCTGCATT	2434
Qy	2330	GAGCATGGGACCTCCGCGACGCGCAAGTCGAGTGCAGCCCTGGCTGGATGCGGAACAC	2389
Db	2435	GAGCAGGGACCTGTAAAGATGCGAAATGTGAATCGCAGAGGGCTGGAATGTGGAACAC	2494
Qy	2390	TGCACCATCGCTCACTATCTCTGGATAGGGTAGTTTAAAGAGGGTTGCCCTGGTGTGCAAT	2449
Db	2495	TGCACCATTGAT-----GGCTGCCCTGATTTGTGCAAC	2527
Qy	2450	GGCAACCGCAGATGTACTTTAGACCTGAATGTGTTGGCACTGCGTCTGCCAGCTGGGCTGG	2509
Db	2528	GGTAACGGGAGATGCAACCTGGGTCAAGACAGCTGGCAGTGTGTCTGTGCCAGACCGGCTGG	2587
Qy	2510	AGAGGAGCTCGCTGTGACACTTCCATGGAGACTGCCTCGCGTGCACAGCAAGAACAATGAT	2569
Db	2588	AGAGGGCTCGAATGCAACAGTTTGCATTGGAAACCTCTCGCTGATACAGGATATAGAG	2847
Qy	2570	GGAGATGGCCTGTGGACTGCATGAGACCCCTGACTGTGCTCCAGGCCCTGTGCCATATC	2629
Db	2648	GGAGATGGCCTGTGGACTGCCTGGACCCCTGACTGTGCTCCCTACAGTCAGCCTGTCAAGAC	2707
Qy	2630	AACCCGCTGTCCTTGGCTCCCTTAACCTCTGGACATCATCCAGGAGACACAGGTCCT	2689
Db	2708	AGCCTGCTCTGCGGGGGTCTCGGAGCCCTTGGACATCATTCAGCAAGTCTCAGACAGAC	2767
Qy	2690	GTGTCAAGCAGAACCTTACACTCCTTCTATGACCGCATCAAGTTCCTCGTGGGCAAGGAC	2749
Db	2768	TGGCCTGCAGTA-----AGTCTTCTATGACCGCATCAAGCTCTTGGCAGCAAGGAC	2821
Qy	2750	AGCAGCAGATATCCCGGGGAGAACCCCTTTGATGGAGGGCATGCTGTGTATTTCGT	2809
Db	2822	AGCACCCACATCATCTCTGGAGACACCCCTTCAATAGCAGCCTGTGTCTCTGATCCGA	2881
Qy	2810	GGCAAGTGTATGATCAGATGGAACCCGCCCTGGTGTGTGAAATCACTTTGTGCAAT	2869
Db	2882	GGCCAGTAGTAACCATGGATGGGACTCCCTTGGTGGGTGTGAATGTGCTTTGTCAAG	2941
Qy	2870	AACCCCTCTTTTGGATATACAATCAGCAGCAAGATGGGAGCTTTGACTTTGGTGAACAAT	2929
Db	2942	TACCCAAAATATGGCTACACCATCACTCGGCCAGGATGGCAGTTTGCCTGATTGGCAAT	3001
Qy	2930	GGGGGCATCTCCATCATCTGCGGTTTCGAGGGGCACTTTTCATCACAGGAGCACACC	2989
Db	3002	GGGGGTTCTGCCCTTGACTCTTCACTTTGACGGAGCCCCCTTTTCAITAGCCAGGAGCGACA	3061
Qy	2990	CTGTGGCTGCATGGGATCGCTTTCTTTGTCAATGGAACCATCATCATGACATGAGGAG	3049
Db	3062	GTGTGGCTGCAATGGAACAGCTTCTATGGCATGGAACAGCAACCTGGTATATGAAGACGAGGAA	3121

QY 3050 AATGAGATTCACAGCTGTGACCTGAGCAATTTTGGCCGCCCCCAACCCAGTCGTCTCTCCA 3109
 Db 3122 AACTCCATCCCAAGCTGTGACCTCAGTGGCTTTTGGCCGCCAGATCAATCATCATCTCC 3181
 QY 3110 TCCCACTGAGCTCTTGGCCAGCTCTGTGACAGAGAAAGGCCCAATGTGCGGAAAT 3169
 Db 3182 TCTCCTGTGCCACCTTCTTACGGCTTCCCTGCTCGAAGCCCAATGTGCGTGAACC 3241
 QY 3170 CAGGCTTTGAGGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGC 3229
 Db 3242 CAGGCTTCTATGAGAAATGAGCTCTCTGTTACCAATGTGAAGCTCCGTTATCTCAGC 3301
 QY 3230 AGCCGAGCCCTGGCTCAAAATCTGTCTGAGGATTCAGCTCACCCACCCAGCAATCCCC 3289
 Db 3302 TCTAGAACTGACGGGTAAAGTCCCTGCTGAAAGATCAACATGACGCACTCCACAGTGCCC 3361
 QY 3290 TTAACCTCATGAAGGTGCACTCATGTAGCGGTGAGGCGCCCTCTTTCAGAAAGTGG 3349
 Db 3362 TTGAACCTCATCAGGTTCACTGATGTTGCTGTAGAGGGGATCTCTTTCAGAAAGTCA 3421
 QY 3350 TTCCCTGAGCCCGCAGACCTGTCTCTATTTATTTTCAATTTGGGACAGACAGCTCTACAAC 3409
 Db 3422 TTCCAGGCTTCTCCCACTAGCCTACACATTCATCTGGGACAGACAGATGCTTATGGC 3481
 QY 3410 CAGAAGGTGTTGGGCTTTCAGAGCCTTTGTTTCCGTGGGTATGATGATGATCTCTGC 3469
 Db 3482 CAAAGGCTTATGGCCCTATCGATGCTGTGTGTCTTGGTGTGAATGAGACCTGC 3541
 QY 3470 CCAGATCTAATCTGTGGGAAAAAGAACACAGTGTGAGGCTATGAAATGACGCG 3529
 Db 3542 CCCAGTCTCATCTGTGGGAAAAAGGACAGCCCTGCTTCAAGGATTCGAGCTGGACCT 3601
 QY 3530 TCCAAGCTTGAGGATGAGGCTAGACAAACATCATGCTCCCTCAACATTCAGAGTGTATC 3589
 Db 3602 TCCAACCTTGAGGCTGGTCCCTGGACAAACACACACCCCTCAATGTGAAAGCGGAATA 3661
 QY 3590 CTGCACAAAGGAATGGGGAACCCAGTTTGTCTCAGAGCCTCTCTGATTTGGGAGC 3649
 Db 3662 CTACACAAAGGAACAGGGAGAACAGTCTCTGACCCAGCAGCTGCCATCATCAGAGC 3721
 QY 3650 ATCATGGGCAATGGGCGGAGAGCAATCTCTGCCCCAGCTGCAACAGCCCTGTGTGAC 3709
 Db 3722 ATCATGGGCAACGGTCGCGCAGAGCAATCTCTGTGCCAGCTGCAATGGCTGTGTGAA 3781
 QY 3710 GGCACAGCTCTGGCCAGTGGCCCTCACCTGTGCTCTGAGGAGCCCTCTATGTG 3769
 Db 3782 GGCACAAACTGTTAGCCCCCTGTGGCCCTGGCTGTGGGATTCGATGGAGCCCTTTGTT 3841
 QY 3770 GGTGATTTCAACTACATTAAGAGATCTTCCCTCTGGAATGTCAACCAATCTCTAGAG 3829
 Db 3842 GGTGATTTCAACTATATCCGCGCATCTTCCCTCTCGAATGTGACCAATCTTTGAG 3901
 QY 3830 CTGAGGAATAAGATTTTCAACATAGTCAAGTCCAGACACAAATACCTACCTGGCCACA 3889
 Db 3902 TTACGAAATAAGAGTTTAAACATAGCAACAGCCCGCAGACACAAATGTAATCTGGCTGTG 3961
 QY 3890 GACCCCATGAGTGGGCGCTCTCTCTTCTGACAGCAACAGCCGCGGGTCTTTAAATC 3949
 Db 3962 GACCCCGTACTGACTACTCTACTGCTCTGACACCAACAGTCCCGAATCTACCGATC 4021
 QY 3950 AAGTCCACTGTGGGTGAAGCACTTGTCAAGAACTCTGAGGTGTTGGGGAACAGGT 4009
 Db 4022 AAGTCTCTGAGCGGAGCAAGACCTGGCTGGAATTCGGAATTTGGCAGGAGCTGGC 4081
 QY 4010 GACAGTGCCTCCCTTTGATGACATCTGCTGCGGGATGTTGGGAGGCGCACAGAGCC 4069
 Db 4082 GAACAATGCTACCCCTTTGATGAAGCCGCTGTGGGATGAGGGAAGGCTGTGGAGCC 4141
 QY 4070 ACACCTCAGCAATCCAGGGGTATACAGTGTGACAAAGTTTGGGCTGATCTACTCTGTGGAT 4129
 Db 4142 ACCCTGATGAGCCCGCAGAGGTATTCAGTAGACAAAGATGGCTTATGTTGTTGAT 4201

QY 4130 GGCACATGATCAGACGATCGATCAGAAATGGGATCATCTCCACCCCTGCTGGCTCTAAT 4189
 Db 4202 GCCACCATGATCCGGAAGGTGGACCAAAACGGAATCATCTCCACCCCTGCTGGGCTCCAAT 4261
 QY 4190 GATCTCATCATCAGCCCGGCCCTCAGCTGTGATTTCTGTCATGGATATTTCCAGGTAAAGA 4249
 Db 4262 GACCTTCAGCTGTCCGACCACTGAGCTGTGATCTGACATGGAGCTGGCCCGAGTCCGT 4321
 QY 4250 CTGGAGTGGCCACACAGACTTACCCATCAACCCCAATGGCAACTACTATTATGTCTCGAC 4309
 Db 4322 CTAGAATGGCCGACAGACCTGCGCGTCAACCCCATGGCAACTCTCTCTGATCGTTCTGGAG 4381
 QY 4310 ACAATGTGGTCTCTCAATCTCTGAAACACACAGGTGCGCATTTGTCGCGGAGGCC 4369
 Db 4382 AACAACGTCATCTCTGGATCAGGAGAACACACAGGTGAGCATCATCGCGGAGCGGCT 4441
 QY 4370 ATGCACTGCCAGGTCCCTGGCATTCACCACTTCTGCTGAAGCAAGTGGCCATCCACGCA 4429
 Db 4442 ATGCACTGCCAGGTTCCCGGCATCGACTACT---CGCTCAGCAAACTCGCCATCCACTCT 4498
 QY 4430 ACCCTGGAGTACGCCACGCTTTGCTGTTTCAACAATGGGCTCTGTATATTTGCTGAG 4489
 Db 4499 GCGCTGGAATCAGCCAGCGCCATTTGCCATTTCTCACTGGGGTCTCTACATCACTGAG 4558
 QY 4490 ACTGATGAGAAAAAGATCAACCGCATCAGGAGGTCAACACTAGTGGAGAGATCTCACTC 4549
 Db 4559 AGGAGGAGAGAGAGATCAACCGCTACGCCAGTCAACCAATGGAGAGATCTGCCCTC 4618
 QY 4550 GTTGTGGGGCCCCCAGTGGCTGTGACTGTAAAAATGATGCAACTGTGATTTGTTTCT 4609
 Db 4619 TTAGCCGGGGCCCTCAGACTGTGACTGCAAAAAACGATGTCAACTGCTGTCTACTCG 4678
 QY 4610 GGAGAGATGTTTATGCCAGGATCAAGTTAAATACCCCATCTTCTTGGCTGTGTGT 4669
 Db 4679 GGAGATGAGCTTTAGCCAGGAGCCCTCTGTAATCTGCGCTCTCTTCTAGCCGTGGCT 4738
 QY 4670 GCTGATGGGAGCTCTACGTGGCGACCTTGGGAAACATCCGAAATTCGGTTTATCCGGAG 4729
 Db 4739 CCGGATGGCACTCTACATTCAGACCTTGGGAATATCCGATCAGGCGGTGAGCAAA 4798
 QY 4730 AACAGCCCTTCTCAACACCCAGAACATGATGAGCTGTCTTCAACATTCACCAAGAG 4789
 Db 4799 AATAAACCCGTTCTTAAACGATTCACCCAGTATGAGGTGCACTCCCGGAGAAACAGGAA 4858
 QY 4790 CTCTATCTCTTTGATACCAACCCGCAAGCACTGTACACCAAGCTTCCCAAGAGAGAC 4849
 Db 4859 TTGTACGTTTCAACGCTGATGTTATCCATCAGTACATCTGTGAGTCTGTGACTGGGAG 4918
 QY 4850 TACCTGTACAACTTCACTTACACTGGGGAGCGGCAATCACTCATCAACAGCAACAAAT 4909
 Db 4919 TACTTGTACAAATTTACATACAGCGCTGACAAATGACGTCAACGAGTTGATTTGACAAAC 4978
 QY 4910 GGCACATGTTAAATGTCGCGGAGACTCTACTGGGATGCCCTCTGCGTGTGTGCTCCA 4969
 Db 4979 GGGATTTCCCTTAAAGATCCCGCGGAGACAGTGGATGGATGCCCGCACTGTCTATGCCG 5038
 QY 4970 GATGGCAGGTACTGGGTGACCATGGGCACCAACAGTGCATCAAGAGTGTGACCAACA 5029
 Db 5039 GATAATCAGATTATCACCTTACTGTGGGCACCAATGGAGGCTCAAAAGCCGTGCTCACT 5098
 QY 5030 CAAGGACAGTGGCCATGATGATACATGCAATTCGGGCTTCTGGCAACCAAA 5089
 Db 5099 CAGAACCTGGAGTGGGCTCATGACTTATGATGGAACACTGGGACTCTCTAGCCCAAG 5158
 QY 5090 AGCAATGAAAACGGATGGACAAATTTATGATGACGACAGCTTTGGCCGCTGACAAAT 5149
 Db 5159 AGTGTGAAAACGGATGGACAACTTTTATGATGATGACCAAGGCGCTGTGACCAAT 5218
 QY 5150 GTGACCTTCCCTACTGGCCAGGTGAGCTTTCCGAAATGATGATGACAGAGTTCAGTGTGAT 5209
 Db 5219 GTGACCCGCCCAACCGGCGTGTGTGACCACTGTGACCCGGGAAATGGAGAAATCTATCAC 5278
 QY 5210 GTCCAGGTAGACCTCCAGCAAGGATGAT---GTCACCAATAACCAACCACTGTCTGCC 5266

Db 5279 ATTGACATTGAGAACTCCAAACCGGATGATGACGTCACTGTGATCAACCAACCTCTCTCC 5338
 QY 5267 TCAGCGGCTTCTACACACTGCTCGAAGCAAACTCGGAAACAGCTACTACATCGGGCC 5326
 Db 5339 GTGGAGGCTCTCTATACAGTGGTACAGATCAAGTGGAAACAGCTACCACTCTGCAAT 5398
 QY 5327 GATGCTCTTTCGGGCTGCTGCTGGCAACCGCATAGAGTGGCGCTGCAAGATGAGCC 5386
 Db 5399 AATGGAACCTCGCGGTGATGTAGCCAAACGCAATGCTGTGACGTTCCACAGTGAGCC 5458
 QY 5387 CACTTGTGGTGGCAGCGTCAACCCACCGTGGCAAGAGGATGTCAGCTGCCATC 5446
 Db 5459 CAGTCTTCGAGGACCATCAACCCACCATCGGGCTGCAACATCTCTCTGCCATG 5518
 QY 5447 GACAAACGCTCAACCTGGTGGAGTGGCGGCGAGCGCAAGAGAGGCTCGGGCCAGGTC 5506
 Db 5519 GAGAATGGCTGAACCTCCATCGAGTGGCGCTGAGGAAGGAACAGATCAAAAGGCAAGTC 5578
 QY 5507 ACTGCTTTGGGCGCGCTCGGGTGCACAAACCGAATCTCTATCTCTGGACTTGTAT 5566
 Db 5579 ACCATCTTTGGAGGAAGCTTCGGGTCACGGAAGGAATCTCTGTCCATTGATTATGAC 5638
 QY 5567 CGCTAACACACACAGAGAAGATCTATGATGACCAACCGCAAGTTCAACCTTCGGATTCTG 5626
 Db 5639 CGAAATATCCGTACGGAGAGATCTACGATGACCAACCGGAATTCACCTTGAGATCATC 5698
 QY 5627 TACGACAGGCGGGCGGCCAGCTCTGTGTCACCCAGAGAGGCTGAATGTGTCAAC 5686
 Db 5699 TATGACAGTGGCGGCCCTCTCTGTGCTCCGAGCAGTGGGCTGGCAGCGCTCAAT 5758
 QY 5687 GTGACATACTCCCTGGGGTTTACATTGCTGGATCCAGAGGGGATCATGCTGAAAGA 5746
 Db 5759 GTCTCTACTTCTCAATGGCGCTTGGCGGCTCCAGAGGGGCGCATGACGAGAGG 5818
 QY 5747 ATGGAATACGACAGCGGGCGGCGATCACATCCAGGATCTTCGCTGATGGGAAGATGG 5806
 Db 5819 ACAGACATGACAAAGAGCGCGATCGTGTCCGCGATGTCGCCAGCGGAAAGTCTGG 5878
 QY 5807 AGCTTACATATCTTAGAAGTCCATGCTGCTACTACACAGCGAGGAGAGTATATC 5866
 Db 5879 AGTTATCTCTATCTTGACAGTCCATGCTCTCTGCTCAGAGCAACCTCAGTACATA 5938
 QY 5867 TTTGAGTTCGACAAAGATGACCGCTCTCTTGTGACGATGCCAACGTTGGCGCGGAG 5926
 Db 5939 TTTGAATATGACTCTCCGATCGCTCCACGCGATCACTATGCCAGTGTCCGCGGAC 5998
 QY 5927 ACATAGAGACCATCCCTAGTGGGTACTACAGAAACATCTATCAGCCCTGAGGGC 5986
 Db 5999 AGCATGCCAGCACCTCCATGTTTACATCCGAAACATTTACAAACCCACCCGAAAGC 6058
 QY 5987 AATGCTCAGTATACAGGACTTCACTGAGGATGGGACCTCTCTTCAACACTTCTACCTG 6046
 Db 6059 AATGATCGGTCACTTTGACTACAGTATGACGGCGCATCTTAAGACATCTTCTTG 6118
 QY 6047 GGCACTGGCGAGGTGATATACAGTATGGCAAACTGTCAAGCTGGCAGAGCGCTC 6106
 Db 6119 GGCACTGGCGAGGTGTTCTCAAGATATGGAATCTCTCAAGTATCAGAGATAGTC 6178
 QY 6107 TATGACACCAAGGTGAGTTTCACTATGACGAGCGGAGGATGCTGAAGACCATC 6166
 Db 6179 TACGACAGCAGCGGTCACTTTGGGTATGACGAGACCAACCGTGTCTGAGATGGTC 6238
 QY 6167 AACCTACAGAAAGAGGCTTCACTGCAACCATCCGCTACCGTACAGATTGGGCGCTGATT 6226
 Db 6239 AATCTCAAAGTGGGCTTCTCTGTACCATCAGGTACCGAAAGGTTGGGCGCTTGTG 6298
 QY 6227 GACGACAGATCTTCGCTTCACTAGGAGGATGTCACCGCGTTTTCATACAC 6286
 Db 6299 GACAAGCAGATTTACAGTTCTCTGAGGAAGGAATGATCAACCGCAGGTTTGAATTATCC 6358
 QY 6287 TAT---GACAAACGCTTCGGGTGACCAAGATGAGGCTGTGATCAACGAGACCCCACTG 6343
 Db

Db 6359 TATACGACAAATAGGTTCCGCAATGCCAGCATCAAAACCGTCATTTAGCGAGACTCCCTT 6418
 QY 6344 CCATTGATCTCTATCGCTATGATGTGAGCAAGACAGAGCAGTTTGGGAAGTTT 6403
 Db 6419 CTGTTGACCTCTACCGTATGACGAGATTTCCGCAAGGTGGACACTTCGGCAAGTTT 6478
 QY 6404 GGTGTCATTTACTATGACATTAACCCAGATCATCAACAGCTGTCTATGACCAACCAAG 6463
 Db 6479 GGGGTCACTTACTAGACATCAACCCAGATCATCAACCTGCGCTCATGACGCTTAGCAAG 6538
 QY 6464 CATTGTTGATGATATGGCAGGATGAAGAGTGCAGATGATGATCTTCCGCTCGCTCATG 6523
 Db 6539 CACTTTGACACCCATGGCGCATCAAGGAAGTGCATATGAGATGTTCCGGTCCCTCATG 6598
 QY 6524 TACTCGATGACCGTCCAGTATGATAACATGGGGGAGTGTGAAGAGGAGCTGAAGTA 6583
 Db 6599 TACTGATGACTGTGCAATATGACAGTATGGTAGGTCATCAAGAGGAACTGAAACTA 6658
 QY 6584 GGACCTACGCAATACCACTCGCTACTCTCTATGATGATGATGCTGACGGCCAGCTGAC 6643
 Db 6659 GGGCCCTATGCCAACACCAAAAGTACACCTATGACTATGACGGGAGCGGCACTCCAG 6718
 QY 6644 ACAGTCTCCATCAATGACAAAGCCACTCTGGGCGCTACAGCTACGACCTCAATGGAACTG 6703
 Db 6719 AGTGTGGCGTCAATGACCGGCTTACCTGGCGCTATAGCTATGACTCAATGGGAACCTG 6778
 QY 6704 CACTTACTGAGCCCTGGGAAACAGTGCAGCGTCAACAACCTACCGTATGATGATCGGAC 6763
 Db 6779 CACTTCTAAACCCAGGAAACAGTGTGCGCTTACGCTTACGCTATGACTCCTCGTAC 6838
 QY 6764 CGCATCACTCGGCTGGGTGAGTGCATCAACAAAGTGGATGAGGATGCTTCTCTGAGCAG 6823
 Db 6839 CGGATACACGCTTAGGGAGCTGCGATGACAAATCGATGAGATGCTTATTTGTCAG 6898
 QY 6824 CGGGCGGTGATATCTTTGAGTACAACTCAGCTGGCTGTCTATCAAGGCTTACAAACGG 6883
 Db 6899 AGAGGTGAGACATCTTTGAAATCAACTCCAAAGGCGCTTCTGACGAGAGCATACAACAG 6958
 QY 6884 GCTGCACTGAGGTGTCAGTACCGTACGATCGCTGGGCGCGCTGCTCCAGCAAG 6943
 Db 6959 GCCAGGATGGAGGTGCGATCCGCTATGACGAGTGGGCGCGGCTTCTTCTACAAG 7018
 QY 6944 AGCAGCCACGCCACCTGCGAGTCTTCTTATGACAGCTGACCAACCCACCAAGGTC 7003
 Db 7019 ACCAAGCTGGGCCACCTACAGTCTTCTTACTCGACCTCCACAAACCCACACGATATC 7078
 QY 7004 ACCAAGCTGTAACCACTCCAGCTCTGAGTACCTCCCTCTACTAGCTTGCAGGA 7063
 Db 7079 ACCCATGTTTACAAACCACTCCAACTCAGATCACTCGCTCTACTATGACCTCCAGGC 7138
 QY 7064 CACTCTTTGCGCATGAGGCTGAGCAGTGGTGTGATGAGTTTACATAGCTTGTGACAACTC 7123
 Db 7139 CACTATTTGCGATGGAGAGCAGTAGTGGTGAAGAACTACTATGCTGCTCAGACAAACAG 7198
 QY 7124 GGGACCCCTCTTGTGCTTTTGTGGAACAGGTTTGTATGATCAAGCAAAATCTCTGTACA 7183
 Db 7199 GGGACCCCTCTTGGCTGTGATCAGTATCAATGGGCTCATGATCAAGCAACTGCAATACACA 7258
 QY 7184 GCCTATGGGAGATCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7243
 Db 7259 GCCTATGGGAGATCTACTATGATCTCCATCCAGACTTCCAGATGCTATGCTTCCAC 7318
 QY 7244 GGTGCGCTTATGATCCACTCAACAGCTTGTCCATGCGGCGGCGAGATTTATGATG 7303
 Db 7319 GGGGCTCTATGATCCCGCTCACCAAGCTGTCTTACTTACTCAACGCTGATGATGATGATG 7378
 QY 7304 CTGCGCGAGCTGAGTACCGGAGCAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 7363
 Db 7379 CTGCGAGAGCTGAGGAGCTCCCGGACTACCAATGAGGAGGAGGAGGAGGAGGAGGAG 7435
 QY 7364 GTCATGCTTTTAACTCTCTATATGTTTCAAAAAACAAACCCCACTGACCAACTCCCCAGG 7423
 Db 7436 CCAGCCCTTCAACCTGTACATGTTTCAAGAAACAAACAACTCTCTGAGCAATGAGCTGAC 7495

Search completed: June 24, 2004, 19:41:25
Job time : 2272 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 16:21:39 ; Search time 81 Seconds
(without alignments)
9650.944 Million cell updates/sec

Title: US-10-029-020-14

Perfect score: 14887

Sequence: 1 MDVKERKPYRSLTRRRDAER.....ELSDSANNIFHRQSEMGRR 2769

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1163542 seqs, 282313646 residues

Total number of hits satisfying chosen parameters: 1163542

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14887	100.0	2769	12	US-10-383-201-44
2	14887	100.0	2769	12	US-10-029-020-14
3	14610	98.1	2775	12	US-10-383-201-56
4	14529	97.6	2771	9	US-09-808-602-82
5	14529	97.6	2771	10	US-09-808-198-70
6	14406.5	96.8	2794	12	US-10-042-865-2
7	10227	68.7	2715	12	US-10-042-865-52
8	10227	68.7	2715	12	US-10-029-020-51
9	10221	68.5	2721	16	US-10-038-854-38
10	10201	68.5	2721	16	US-10-038-854-36
11	9888.5	66.4	2628	16	US-10-038-854-40
12	9856	66.2	2613	16	US-10-038-854-42
13	9616.5	64.6	2802	9	US-09-808-602-81
14	9616.5	64.6	2802	10	US-09-808-198-69
15	9616.5	64.6	2802	12	US-10-072-012-489

16	9610.5	64.6	2764	9	US-09-808-602-80
17	9610.5	64.6	2764	10	US-09-808-198-68
18	9610.5	64.6	2764	12	US-10-072-012-487
19	9601	64.5	2765	9	US-09-808-602-84
20	9601	64.5	2765	10	US-09-808-198-72
21	9601	64.5	2765	12	US-10-072-012-488
22	9584	64.4	2725	12	US-10-029-020-52
23	9584	64.4	2725	16	US-10-408-765A-1687
24	9536	64.1	2733	9	US-09-808-602-8
25	9536	64.1	2733	10	US-09-808-198-8
26	9530.5	64.0	2724	9	US-09-808-602-13
27	9530.5	64.0	2724	10	US-09-808-198-13
28	9501	63.8	2725	15	US-10-295-027-928
29	9418	63.3	2759	12	US-10-072-012-144
30	9314	62.6	2633	15	US-10-144-194A-82
31	9303	62.5	2590	12	US-10-072-012-490
32	9145	61.4	2346	12	US-10-072-012-491
33	8874	59.6	2551	15	US-10-144-194A-80
34	6389.5	42.9	1737	9	US-09-808-602-83
35	6389.5	42.9	1737	10	US-09-808-198-71
36	6258.5	42.0	1688	15	US-10-144-194A-113
37	3979.5	26.7	1045	12	US-10-042-865-54
38	3979.5	26.7	1045	12	US-10-029-020-54
39	3854	25.9	730	12	US-10-042-865-51
40	3854	25.9	730	12	US-10-029-020-50
41	3854	25.9	730	16	US-10-408-765A-2505
42	3804.5	25.6	2515	12	US-10-042-865-53
43	3804.5	25.6	2515	12	US-10-029-020-53
44	3730	25.1	964	12	US-10-042-865-55
45	2989	20.1	768	9	US-09-773-517-11

ALIGNMENTS

RESULT 1

US-10-383-201-44
; Sequence 44, Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHO
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 44
; LENGTH: 2769
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-383-201-44

Query Match		100.0%	Score 14887	DB 12	Length 2769		
Best Local Similarity		100.0%	Pred No. 0				
Matches 2769		Conservative 0	Mismatches 0	Indels 0	Gaps 0		
Qy	1	MDVKERKPYSLTRRRDAERYTSSADSEBGAPOKSYSSSTLTKAYDQDARLAYGSRV	60				
Db	1	MDVKERKPYSLTRRRDAERYTSSADSEBGAPOKSYSSSTLTKAYDQDARLAYGSRV	60				
Qy	61	KDIVPQAEBCFCRGANFTREIGLEBVTTPHGTLYRTDGLPQCGYSWGAGSDADMEAD	120				
Db	61	KDIVPQAEBCFCRGANFTREIGLEBVTTPHGTLYRTDGLPQCGYSWGAGSDADMEAD	120				
Qy	121	TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHENTETDHPGGLQNHARLTRPP	180				
Db	121	TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHENTETDHPGGLQNHARLTRPP	180				
Qy	181	PPLSHAHTPNQHAASINSINRGNTFRSPNPSAPTDSLSGEPAGGAQEPHAQENWL	240				
Db	181	PPLSHAHTPNQHAASINSINRGNTFRSPNPSAPTDSLSGEPAGGAQEPHAQENWL	240				
Qy	241	LNSNIPLETNGLKQPLGLTQDNLTEMDLIGASRHDGAYS DGHFLFKPGGTSPLEFCTTS	300				
Db	241	LNSNIPLETNGLKQPLGLTQDNLTEMDLIGASRHDGAYS DGHFLFKPGGTSPLEFCTTS	300				
Qy	301	PGYPLTSTSYSPPPRLPRSTFARPAFNLKKPSKYCNWKCAALSIVAISATLIVILLAYF	360				
Db	301	PGYPLTSTSYSPPPRLPRSTFARPAFNLKKPSKYCNWKCAALSIVAISATLIVILLAYF	360				
Qy	361	VAMHFLGNLHLOPMQOMYEITTEDASWPVPTDVSPLYPSGGTGLTETDPRKGGTTECK	420				
Db	361	VAMHFLGNLHLOPMQOMYEITTEDASWPVPTDVSPLYPSGGTGLTETDPRKGGTTECK	420				
Qy	421	PSGFFPEDSDISGEIDVGRASQKIPPGTFWRSQVFIIDHPVHLKFNVSGLKAAALVGIYG	480				
Db	421	PSGFFPEDSDISGEIDVGRASQKIPPGTFWRSQVFIIDHPVHLKFNVSGLKAAALVGIYG	480				
Qy	481	RKGLPPSHTOFDVELLDGRLLTQBARSLGTPROSGTVPSSSHETGFIQYLDGSIWH	540				
Db	481	RKGLPPSHTOFDVELLDGRLLTQBARSLGTPROSGTVPSSSHETGFIQYLDGSIWH	540				
Qy	541	LAFVNDGKESEVVSFLTTAIESVDNCPNCGNGDCISGTCFLGLFGLPDGCRASCPLV	600				
Db	541	LAFVNDGKESEVVSFLTTAIESVDNCPNCGNGDCISGTCFLGLFGLPDGCRASCPLV	600				
Qy	601	CSNGQYMKRCJLCHSGWKGAECVPTNQCIDVACSNHGTCTGTGICINPGYKGSCEEV	660				
Db	601	CSNGQYMKRCJLCHSGWKGAECVPTNQCIDVACSNHGTCTGTGICINPGYKGSCEEV	660				
Qy	661	DCMDPTCSGRGVCVRGECHFCVWGCTNCETPRATCLDQCSHGHTPLPDTGLCSDPSWT	720				
Db	661	DCMDPTCSGRGVCVRGECHFCVWGCTNCETPRATCLDQCSHGHTPLPDTGLCSDPSWT	720				
Qy	721	GHDCSIICAADCGGHGVCVGTGTCRCEBGMGAACDQRAHPCABSHGTCRDGKCECSPG	780				
Db	721	GHDCSIICAADCGGHGVCVGTGTCRCEBGMGAACDQRAHPCABSHGTCRDGKCECSPG	780				
Qy	781	WNGEHTCIAHLDRVVKEGCPGLCNGNCTDLNGHWCVCOLGWEGACDTSMETACGD	840				
Db	781	WNGEHTCIAHLDRVVKEGCPGLCNGNCTDLNGHWCVCOLGWEGACDTSMETACGD	840				
Qy	841	SKNDGDGLVDCMDPCCQLPLCHINPLCLGSPNLDIIQETQVPVSQNLHSDYDRIFK	900				
Db	841	SKNDGDGLVDCMDPCCQLPLCHINPLCLGSPNLDIIQETQVPVSQNLHSDYDRIFK	900				
Qy	901	LVRDSTHIIIPGENPDGGHACVIRQVMTSDGTPLVGNISFVNNPLFGYIISQDGSF	960				
Db	901	LVRDSTHIIIPGENPDGGHACVIRQVMTSDGTPLVGNISFVNNPLFGYIISQDGSF	960				
Qy	961	DLVTNGGISIIILFERAPFITQHTLWLPWDRFPFVNMETIIMRHEENIEIPSCDLSNFARP	1020				
Db	961	DLVTNGGISIIILFERAPFITQHTLWLPWDRFPFVNMETIIMRHEENIEIPSCDLSNFARP	1020				
Qy	1021	PVVSPLTSFASCAEKGPVPEIQALQBEISISGCKVRLSYLSRTPGYKSVLRISLT	1080				

1021	PVVSPLTSFASCAEKGPVPEIQALQBEISISGCKVRLSYLSRTPGYKSVLRISLT	1080
1081	HPTTIPFNLMKVHLMVAVEGRLFRKWFAPAAADLSYFIWDTKTDVYNQKVFGLSEAFVSGV	1140
1081	HPTTIPFNLMKVHLMVAVEGRLFRKWFAPAAADLSYFIWDTKTDVYNQKVFGLSEAFVSGV	1140
1141	EYESCPDLILWEKRTTVLQGYEIDASKLGWSLQKXHALNIQSGILHKGNGENQFVSQQP	1200
1141	EYESCPDLILWEKRTTVLQGYEIDASKLGWSLQKXHALNIQSGILHKGNGENQFVSQQP	1200
1201	PVIGSIMGNGRRRSISCPSCNGLADNKLAPVALTCGSDGSLYVGFNVIIRIPPSGNV	1260
1201	PVIGSIMGNGRRRSISCPSCNGLADNKLAPVALTCGSDGSLYVGFNVIIRIPPSGNV	1260
1261	TNILELRNKDPRHSHSPAHKYIYLATDPMGAVFLSDNSRRVFKIKSTVAVKOLVKNSEV	1320
1261	TNILELRNKDPRHSHSPAHKYIYLATDPMGAVFLSDNSRRVFKIKSTVAVKOLVKNSEV	1320
1321	VAGTGDQCLPDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDTGTMIRRIDQNGIIST	1380
1321	VAGTGDQCLPDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDTGTMIRRIDQNGIIST	1380
1381	LLGNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYLDNNVLIQISENHQVRI	1440
1381	LLGNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYLDNNVLIQISENHQVRI	1440
1441	VAGPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIQVTTTS	1500
1441	VAGPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIQVTTTS	1500
1501	GEISILVAGAPSGCDKNDANCDPFGDDGYAKADAKLTPSSLAACADGELYADLGNIRI	1560
1501	GEISILVAGAPSGCDKNDANCDPFGDDGYAKADAKLTPSSLAACADGELYADLGNIRI	1560
1561	RFIRKKNPFLLTQNMWELSSPIQELYLFDITGKHLTQSLPTGDIYLYNFTYTGDDGITL	1620
1561	RFIRKKNPFLLTQNMWELSSPIQELYLFDITGKHLTQSLPTGDIYLYNFTYTGDDGITL	1620
1621	ITDNGNKNVVRDRSTGMPLWLVDPDQVYVVTMTGNTSALKSVTTQGHELAMNTYFNGSG	1680
1621	ITDNGNKNVVRDRSTGMPLWLVDPDQVYVVTMTGNTSALKSVTTQGHELAMNTYFNGSG	1680
1681	LLATKSNENGWTTIEYDVSFGRLLTNVTFPGQVSSPRSDTSSVHVHVOVETSSKDDVTIIT	1740
1681	LLATKSNENGWTTIEYDVSFGRLLTNVTFPGQVSSPRSDTSSVHVHVOVETSSKDDVTIIT	1740
1741	NLSASGAFYTLLOQVNRNSYVIGADGSLRLLLANGMEVALQTEPHLIAGTWNPTVGRNV	1800
1741	NLSASGAFYTLLOQVNRNSYVIGADGSLRLLLANGMEVALQTEPHLIAGTWNPTVGRNV	1800
1801	TLPTDNGNLNVEMHQRKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTKIYDDHRKFT	1860
1801	TLPTDNGNLNVEMHQRKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTKIYDDHRKFT	1860
1861	LRILYDQAGPSLWSPSSRLNGVNVVYSPGYIAGTQRMISERMEYDQAGRITSRIFAD	1920
1861	LRILYDQAGPSLWSPSSRLNGVNVVYSPGYIAGTQRMISERMEYDQAGRITSRIFAD	1920
1921	GKTSVYTLKSMVLLLSHQRYIFEFKNDRLSSVTMPNVARQTLETIRSVGYRNIYQ	1980
1921	GKTSVYTLKSMVLLLSHQRYIFEFKNDRLSSVTMPNVARQTLETIRSVGYRNIYQ	1980
1981	PPENASVIOPTEDGHLHFTYLGTRRVYKYVGLSKLAETLYDTTKVSYFYDETAGM	2040
1981	PPENASVIOPTEDGHLHFTYLGTRRVYKYVGLSKLAETLYDTTKVSYFYDETAGM	2040
2041	LKTINLQNEGTCTIRYRQIGPLIDRQIFRFTTEGMVNAFVDYNDNSFRVTSQVAVINE	2100
2041	LKTINLQNEGTCTIRYRQIGPLIDRQIFRFTTEGMVNAFVDYNDNSFRVTSQVAVINE	2100
2101	TPLPIDRYDDVSGKTEQKFGVYIYDINOIITAVMTHTHFDAYGEMKVOVEIFR	2160

Db 2101 TPLPIDLYRDDVSGKTEQKQKGVYIYDINQIITAVMTHTKHFADYGRMKEVQVEIPR 2160
Qy 2161 SLMYMTVQYDNNMGRVVKELKVPYANTRYSEYEDADQLOQTVSINDKPLWRYSYDLN 2220
Db 2161 SLMYMTVQYDNNMGRVVKELKVPYANTRYSEYEDADQLOQTVSINDKPLWRYSYDLN 2220
Qy 2221 GNLHLSPGNSARLTPLRYDIRITELGDVOVQKMDDEGFLRQSGDIEFYNSAGLLIKA 2280
Db 2221 GNLHLSPGNSARLTPLRYDIRITELGDVOVQKMDDEGFLRQSGDIEFYNSAGLLIKA 2280
Qy 2281 YNRAGSWVRVYDGLGRRVSSKSSHLLQFFYADLTNPTKVTHLYNHSSSEITSLYYD 2340
Db 2281 YNRAGSWVRVYDGLGRRVSSKSSHLLQFFYADLTNPTKVTHLYNHSSSEITSLYYD 2340
Qy 2341 LQHLFAMELSSGDEFYIACDNIGTFLAVFSGTGLMIKQILYTAIGEIVYMDTNPQIIII 2400
Db 2341 LQHLFAMELSSGDEFYIACDNIGTFLAVFSGTGLMIKQILYTAIGEIVYMDTNPQIIII 2400
Qy 2401 GYHGGLYDPLTKLVHNGRRDYDLVAGRWTSFDPHKLWHLSSSNVMPFNLYMFQNNPNISN 2460
Db 2401 GYHGGLYDPLTKLVHNGRRDYDLVAGRWTSFDPHKLWHLSSSNVMPFNLYMFQNNPNISN 2460
Qy 2461 SODIKCFMTDVNSWLLTFFQLHNVIIPGYPKPMDAMEPSYELIHTQMKTQEWDNKSIL 2520
Db 2461 SODIKCFMTDVNSWLLTFFQLHNVIIPGYPKPMDAMEPSYELIHTQMKTQEWDNKSIL 2520
Qy 2521 GVOCEVQKQKAFVTLERPDOLYGSTITSQQAPKTKFASGSGVFGKVKPALKDGRVT 2580
Db 2521 GVOCEVQKQKAFVTLERPDOLYGSTITSQQAPKTKFASGSGVFGKVKPALKDGRVT 2580
Qy 2581 TDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHEYFVKPGPSEGDAILGLSGGRT 2640
Db 2581 TDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHEYFVKPGPSEGDAILGLSGGRT 2640
Qy 2641 LENGVNVTQINTVLNGRRRTTDLQYQALCLNTRYGTTTIDDEKARVLELARQAVR 2700
Db 2641 LENGVNVTQINTVLNGRRRTTDLQYQALCLNTRYGTTTIDDEKARVLELARQAVR 2700
Qy 2701 QAWAREOORLREBEGELRANTEGEKQOVLSTGRVQYDGFVVISVQYPELSDSANNIHF 2760
Db 2701 QAWAREOORLREBEGELRANTEGEKQOVLSTGRVQYDGFVVISVQYPELSDSANNIHF 2760
Qy 2761 MRQSEMGR 2769
Db 2761 MRQSEMGR 2769

RESULT 2

US-10-029-020-14
; Sequence 14, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29

; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 2769
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-029-020-14

Query Match 100.0%; Score 14887; DB 12; Length 2769;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDVKKERKPYRSUTRRDAERRYTSSADSESGKAPQKSYSSSETLKAYDQDARLAYGSRV 60
Db 1 MDVKKERKPYRSUTRRDAERRYTSSADSESGKAPQKSYSSSETLKAYDQDARLAYGSRV 60
Qy 61 KDIVPOEAEEFCRTGANFTLRELGLVEVTPPHGTLYRTDGLPQCGYSMGAGSDADMEAD 120
Db 61 KDIVPOEAEEFCRTGANFTLRELGLVEVTPPHGTLYRTDGLPQCGYSMGAGSDADMEAD 120
Qy 121 TVLSPHVRVLWGRSTRSGRSSCLSSRANSNLTTTDEHEHENTETDHPGCLQNHARLATPP 180
Db 121 TVLSPHVRVLWGRSTRSGRSSCLSSRANSNLTTTDEHEHENTETDHPGCLQNHARLATPP 180
Qy 181 PPLSHAHTPNQHHAAASINSLARNGFTPRSNPSPAPTDHSLSGEPAGGAQPAHAQENWL 240
Db 181 PPLSHAHTPNQHHAAASINSLARNGFTPRSNPSPAPTDHSLSGEPAGGAQPAHAQENWL 240
Qy 241 LNSNIPLSTRNLKQPFLLGTQDNLIEMDIILGASRHGAYSDGHFLFKPGGTSPLFCTTS 300
Db 241 LNSNIPLSTRNLKQPFLLGTQDNLIEMDIILGASRHGAYSDGHFLFKPGGTSPLFCTTS 300
Qy 301 PGYPLTSTVTSVPPRPLPRSTFAPAPNLKPKSKYCNWKAALSAIVISATLVILLAYF 360
Db 301 PGYPLTSTVTSVPPRPLPRSTFAPAPNLKPKSKYCNWKAALSAIVISATLVILLAYF 360
Qy 361 VAMHLFGLNWHLQPMEGQMYEITDASSVPVTDVSLYPSGGTGLETPDRKGGTTEGK 420
Db 361 VAMHLFGLNWHLQPMEGQMYEITDASSVPVTDVSLYPSGGTGLETPDRKGGTTEGK 420
Qy 421 PSSFPFDSFIDSGEIDVGRASQKIPGTFWRQVFDHPVHLKFNKFNKAAALVGIY 480
Db 421 PSSFPFDSFIDSGEIDVGRASQKIPGTFWRQVFDHPVHLKFNKFNKAAALVGIY 480
Qy 481 RKGLPPSHQTFDFVLLDGRLLTQEARSLGTPRQSRGTVPSPSHETGFTQYLDGSIWH 540
Db 481 RKGLPPSHQTFDFVLLDGRLLTQEARSLGTPRQSRGTVPSPSHETGFTQYLDGSIWH 540
Qy 541 LAFYNDGKSEVVSFLTTAIESVDNCPNVCYNGNDCISGTCFGLGFLGPGDCGRASCPVL 600
Db 541 LAFYNDGKSEVVSFLTTAIESVDNCPNVCYNGNDCISGTCFGLGFLGPGDCGRASCPVL 600
Qy 601 CSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCTITGTCINPGYKGSCEV 660
Db 601 CSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCTITGTCINPGYKGSCEV 660
Qy 661 DCMDPTCSGRGVCVRGECHCFVGMGGTNCETPRATCLDQCSCGHGTFLLPDTGLCSODPSWT 720
Db 661 DCMDPTCSGRGVCVRGECHCFVGMGGTNCETPRATCLDQCSCGHGTFLLPDTGLCSODPSWT 720
Qy 721 GHDCSIEICADCGGHGVCVGGTCBCEGWNWGAACDQACHPRCAEHGTCRDGKCECPG 780
Db 721 GHDCSIEICADCGGHGVCVGGTCBCEGWNWGAACDQACHPRCAEHGTCRDGKCECPG 780
Qy 781 WNGEHCTTAHYLDRVVVKEGCPCLNGNRCCTLDLNGHCHVCQOLGWRGAGCDDTSMETACGD 840
Db 781 WNGEHCTTAHYLDRVVVKEGCPCLNGNRCCTLDLNGHCHVCQOLGWRGAGCDDTSMETACGD 840
Qy 841 SKDNDGDGLVDCMDPCCCLQPLCHINPLCLGSPNPLDIIQETQVPVVSQNLHFSFYDRKIF 900
Db 841 SKDNDGDGLVDCMDPCCCLQPLCHINPLCLGSPNPLDIIQETQVPVVSQNLHFSFYDRKIF 900

QY 901 LVGRDSTHIIIPENPFPGGHACVIRGQVMTSDGTPLVGNISFVNPLFGYTIISRODGSF 960
DB 901 LVGRDSTHIIIPENPFPGGHACVIRGQVMTSDGTPLVGNISFVNPLFGYTIISRODGSF 960
QY 961 DLVNTNGISIIILRRERAPFIIIOBHTLWLPDRFPVMEIIIMRHEBNEIIPSCDLSNFARN 1020
DB 961 DLVNTNGISIIILRRERAPFIIIOBHTLWLPDRFPVMEIIIMRHEBNEIIPSCDLSNFARN 1020
QY 1021 PVVSPSLTSPASSCAEKGPVPEIQALQEBEISGCKMRLSYLSRTPGYKSVLRISLT 1080
DB 1021 PVVSPSLTSPASSCAEKGPVPEIQALQEBEISGCKMRLSYLSRTPGYKSVLRISLT 1080
QY 1081 HPTIPFNLMKVLHMAVEGBLFRKWFAPAAPDLISYFIWKTVDYNOKVFLGSEAFVSVGY 1140
DB 1081 HPTIPFNLMKVLHMAVEGBLFRKWFAPAAPDLISYFIWKTVDYNOKVFLGSEAFVSVGY 1140
QY 1141 EYESCPDLILWEKRTTVLQGEYIDASKLGWSLDKXHALNIQSGILHKGNGENQFVSQOP 1200
DB 1141 EYESCPDLILWEKRTTVLQGEYIDASKLGWSLDKXHALNIQSGILHKGNGENQFVSQOP 1200
QY 1201 PVIGSINGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVDFNYIRRIIPSGNV 1260
DB 1201 PVIGSINGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVDFNYIRRIIPSGNV 1260
QY 1261 TNILELRNKDPRHSHSPAHKYLYLATDPMGSAVFLSDNSRRVPFKISTVAVKDLVKNSEV 1320
DB 1261 TNILELRNKDPRHSHSPAHKYLYLATDPMGSAVFLSDNSRRVPFKISTVAVKDLVKNSEV 1320
QY 1321 VAGTGDQCLPDDTRCGDGGKATEATLTNPRGTTVDKFGLIYFVDTGMIRRIDQNGIIST 1380
DB 1321 VAGTGDQCLPDDTRCGDGGKATEATLTNPRGTTVDKFGLIYFVDTGMIRRIDQNGIIST 1380
QY 1381 LLGSNDLTSARPLSCDSVMDISQVLEWPTDPLAINPMDSNLYVLDNNVQLQISEKHQVRI 1440
DB 1381 LLGSNDLTSARPLSCDSVMDISQVLEWPTDPLAINPMDSNLYVLDNNVQLQISEKHQVRI 1440
QY 1441 VAGRPMHCQVPGIDHFLLSKVAIHALESATALAVSHNGVLYIAETDEKKINRIQVITS 1500
DB 1441 VAGRPMHCQVPGIDHFLLSKVAIHALESATALAVSHNGVLYIAETDEKKINRIQVITS 1500
QY 1501 GEISLVAGAPSGCDCKNDANCDCFSGDGYAKDAKNTSPSSLAVCADGELYVADLGNTRI 1560
DB 1501 GEISLVAGAPSGCDCKNDANCDCFSGDGYAKDAKNTSPSSLAVCADGELYVADLGNTRI 1560
QY 1561 RFIKKNKPFNTQNMVELSPIDQELYLFDTTGKHLYTOSLPDGYLYNFTYTGDDITL 1620
DB 1561 RFIKKNKPFNTQNMVELSPIDQELYLFDTTGKHLYTOSLPDGYLYNFTYTGDDITL 1620
QY 1621 ITDNNGMNVNRDSTGMPILWLVVDPGVVYVMTGNSALKSVTTQGHELAMWYHNSG 1680
DB 1621 ITDNNGMNVNRDSTGMPILWLVVDPGVVYVMTGNSALKSVTTQGHELAMWYHNSG 1680
QY 1681 LLATKSNENGWTFYEDSFGRLTNVPTPGQVSSFSRSTDSVHVQVETSKDDVTIIT 1740
DB 1681 LLATKSNENGWTFYEDSFGRLTNVPTPGQVSSFSRSTDSVHVQVETSKDDVTIIT 1740
QY 1741 NLSASGAFYTLLOQDVNRNSYIIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKENV 1800
DB 1741 NLSASGAFYTLLOQDVNRNSYIIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKENV 1800
QY 1801 TLPIDNGLNLVWRQKQARGQVTVFGRRLRVHNNLSLDFDRVTRTEKIIDHRRKPT 1860
DB 1801 TLPIDNGLNLVWRQKQARGQVTVFGRRLRVHNNLSLDFDRVTRTEKIIDHRRKPT 1860
QY 1861 LRILYQAGRPSSLWSPSSRLNGVNVYSPGVIAGIQRGIMRMEYDOAGRITSIFAD 1920
DB 1861 LRILYQAGRPSSLWSPSSRLNGVNVYSPGVIAGIQRGIMRMEYDOAGRITSIFAD 1920
QY 1921 GKTWSYTYLEKSMVLLHLSORQYIFBFDKXNRLSSVTMPNVARQTLTIRSVGYRYNYQ 1980
DB 1921 GKTWSYTYLEKSMVLLHLSORQYIFBFDKXNRLSSVTMPNVARQTLTIRSVGYRYNYQ 1980

QY 1981 PPEGNASVIOPTDEGHLLHTFYLGTGRRVYIKYKLSKLAETLYDTTKVSFTYDETAGM 2040
DB 1981 PPEGNASVIOPTDEGHLLHTFYLGTGRRVYIKYKLSKLAETLYDTTKVSFTYDETAGM 2040
QY 2041 LKTINLQNEGTCITIRVQIGPLIDROIIFRPTBEGMNARFDYNDNSFRVTSQAVINE 2100
DB 2041 LKTINLQNEGTCITIRVQIGPLIDROIIFRPTBEGMNARFDYNDNSFRVTSQAVINE 2100
QY 2101 TPLPIDLYRYDDVSGKTEQFGKFGVIVYDINQIITTAVMTHTHKHFDAYGRMKEVQYEIFR 2160
DB 2101 TPLPIDLYRYDDVSGKTEQFGKFGVIVYDINQIITTAVMTHTHKHFDAYGRMKEVQYEIFR 2160
QY 2161 SLMTYMTVOYDNMGRRVVKELKVGFPYANTTRYSEYDADQOLQTVSINDKPLWRYSYDLN 2220
DB 2161 SLMTYMTVOYDNMGRRVVKELKVGFPYANTTRYSEYDADQOLQTVSINDKPLWRYSYDLN 2220
QY 2221 GNLHLLSPGNSARLTPLRYDIRDRIITRLGDVQYQWDEDDGFLRQGGDIFPYNSAGLLIKA 2280
DB 2221 GNLHLLSPGNSARLTPLRYDIRDRIITRLGDVQYQWDEDDGFLRQGGDIFPYNSAGLLIKA 2280
QY 2281 YNRAGSWRVRYDGLGRRVSSKSHHLLQFFVADLTNPTKVTHLYNHSSSITSLYYD 2340
DB 2281 YNRAGSWRVRYDGLGRRVSSKSHHLLQFFVADLTNPTKVTHLYNHSSSITSLYYD 2340
QY 2341 LQGHLLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAIGEIVMDTNPFOIIL 2400
DB 2341 LQGHLLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAIGEIVMDTNPFOIIL 2400
QY 2401 GYHGLGYDPLTKLVHMGRRDYDVLAGRWTSPDHLMKHLSSNNVMPFNLYMFKNNPISN 2460
DB 2401 GYHGLGYDPLTKLVHMGRRDYDVLAGRWTSPDHLMKHLSSNNVMPFNLYMFKNNPISN 2460
QY 2461 SODIKCMTDVNSHLLTFGLHNVIIPGYKPKDMDAMEPSYELIHTQMTQEWDNKSIL 2520
DB 2461 SODIKCMTDVNSHLLTFGLHNVIIPGYKPKDMDAMEPSYELIHTQMTQEWDNKSIL 2520
QY 2521 GVQCEVQKQKAFVTLRERFQLYGSTITSQQAPKTKKFASSGSGVFGKVFALKDGRVT 2580
DB 2521 GVQCEVQKQKAFVTLRERFQLYGSTITSQQAPKTKKFASSGSGVFGKVFALKDGRVT 2580
QY 2581 TDIISVANEDGRRVAAIILNHAHYLENLHFTIDGVTHYFVKPGPSBGDLAILGLSGRRT 2640
DB 2581 TDIISVANEDGRRVAAIILNHAHYLENLHFTIDGVTHYFVKPGPSBGDLAILGLSGRRT 2640
QY 2641 LENGVNVTVSQINTVLNGRTRRYTDIOQYCALCLNTRYGTTIDEEKARVLELARQAVR 2700
DB 2641 LENGVNVTVSQINTVLNGRTRRYTDIOQYCALCLNTRYGTTIDEEKARVLELARQAVR 2700
QY 2701 QAWAREOQRLREBEGELRAWTEGKQOVLSTGRVQYDGFVIVSVEQYPELSDSANNIHF 2760
DB 2701 QAWAREOQRLREBEGELRAWTEGKQOVLSTGRVQYDGFVIVSVEQYPELSDSANNIHF 2760
QY 2761 MRQSEMGRR 2769
DB 2761 MRQSEMGRR 2769

RESULT 3

US-10-383-201-56
; Sequence 56: Application US/10383201
; Publication No. US20040029226A1
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-568A
; CURRENT APPLICATION NUMBER: US/10/383,201
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 10/029020
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/365,984
; PRIOR FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/372,022
; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/389,143
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/391,779
; PRIOR FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/410,755
; PRIOR FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: 60/412,957
; PRIOR FILING DATE: 2002-09-23
; PRIOR APPLICATION NUMBER: 10/051,874
; PRIOR FILING DATE: 2002-01-16
; PRIOR APPLICATION NUMBER: 60/366,928
; PRIOR FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 10/055,877
; PRIOR FILING DATE: 2002-01-22
; NUMBER OF SEQ ID NOS: 155
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 56
; LENGTH: 2775
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-383-201-56

Query Match 98.1%; Score 14610; DB 12; Length 2775;
Best Local Similarity 98.2%; Pred No. 0;
Matches 2743; Conservative 2; Mismatches 6; Indels 42; Gaps 12;

Qy 1 MDVKERPYSLTRRRDAERYTSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60
Db 1 MDVKERPYSLTRRRDAERYTSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60

Qy 61 KDIVPOAEBCFRGANFTLREIGLEEVTPPHGTLRYRTDGLPQCGYSMGAGSDADMEAD 120
Db 61 KDIVPOAEBCFRGANFTLREIGLEEVTPPHGTLRYRTDGLPQCGYSMGAGSDADMEAD 120

Qy 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHEHENTETDHPGGLQNHARLRTTP 180
Db 121 TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDEHEHENTETDHPGGLQNHARLRTTP 180

Qy 181 PPLSHATPNQHAASINSLNRGFTPRSNPSPAPTDHLSGEPGAGGAOEPHAQENML 240
Db 181 PPLSHATPNQHAASINSLNRGFTPRSNPSPAPTDHLSGEPGAGGAOEPHAQENML 240

Qy 241 LNSNIPLETNLGKOPFLGTLQDNLIEMDILGASRHDGAYSDDGHLFKPGGTSPLFCTTS 300
Db 241 LNSNIPLETNLGKOPFLGTLQDNLIEMDILGASRHDGAYSDDGHLFKPGGTSPLFCTTS 300

Qy 301 PGVPLTSTVYSPPPRLPRSTFARPAFNLKPKSKYCNWKCAALSALVISATLVIILAYF 360
Db 301 PGVPLTSTVYSPPPRLPRSTFARPAFNLKPKSKYCNWKCAALSALVISATLVIILAYF 360

Qy 361 VAMHPLGLNHLQHPMEQOMYEITEDTASSWPVPTDVSILYPSGTLGLETDRKKGTEGK 420
Db 361 VAMHPLGLNHLQHPMEQOMYEITEDTASSWPVPTDVSILYPSGTLGLETDRKKGTEGK 420

Qy 421 PS9FFPPDSFDGSEIDVGRASQKIPPGTFWRSQVFIIDHPVHLKFNVSILGKAALVGIY 480
Db 421 PS9FFPPDSFDGSEIDVGRASQKIPPGTFWRSQVFIIDHPVHLKFNVSILGKAALVGIY 480

Qy 481 RKLPPSHQTFDFVELLDGRLLLTQEARSLGTPROSGRGVPPSSHETGPIQYLDSGIWH 540
Db 481 RKLPPSHQTFDFVELLDGRLLLTQEARSLGTPROSGRGVPPSSHETGPIQYLDSGIWH 540

Qy 541 LAFVNDKSEVVSFLTATIESVDNCPNSCYNGDCISGTCCHCFLGFLGPDGGRASCPVL 600
Db 541 LAFVNDKSEVVSFLTATIESVDNCPNSCYNGDCISGTCCHCFLGFLGPDGGRASCPVL 600

Qy 601 CSGNGQYMKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTCTITGTCINPGYKGCSEBV 660
Db 599 CSGNGQYMKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTCTITGTCINPGYKGCSEB- 657

Qy 661 DQMDPTCSGRGVCVRGCHGCVGNGGTNCETPRATCLDQCSGHGTFLPDTGLCSDPSWT 720
Db 658 DQMDPTCSGRGVCVRGCHGCVGNGGTNCETPRATCLDQCSGHGTFLPDTGLCSDPSWT 717

721 GHDCSIEICAAACGGHGVCGGTCTCEBGMWGAACDQACHPRCAEHGTCRKGCECSPG 780
718 GHDCSIEICAAACGGHGVCGGTCTCEBGMWGAACDQACHPRCAEHGTCRKGCECSPG 777
781 WNGEHCTI--AHYLDREVVK--EGCPEGLCNGRCRTLDLNGHVCVQQLGWRGAGCCTSMT 836
778 WNGEHCTISLAHYLDRVVKLSEGCFLCNGRCRTLDLNGHVCVQQLGWRGAGCCTSMT 837
837 ACGDSKNDGDGLVDCMPDDCCLQPLCHINPLCLGSPNPDLIIQSTQVPSQONLHSFYD 896
838 ACGDSKNDGDGLVDCMPDDCCLQPLCHINPLCLGSPNPDLIIQSTQVPSQONLHSFYD 897
897 RIKFLVGRDSTHIIPEGNPFDDGHCACVIRGOVMTSDGTPLVGVNISFVNNPLFGVTISRO 956
898 RIKFLVGRDSTHIIPEGNPFDDGHCACVIRGOVMTSDGTPLVGVNISFVNNPLFGVTISRO 957
957 DGSFDLVNGGSIILRPERAPFITQEHNTLMLPWRDFFVMTETIMRHEENBIIPSCDLSNF 1016
958 DGSFDLVNGGSIILRPERAPFITQEHNTLMLPWRDFFVMTETIMRHEENBIIPSCDLSNF 1017
1017 ASPNPVPSPSPLTSFASCAEKGPVPEIQALQOEISISGCKMRLSYLSSTTPGVKSVLR 1076
1018 ASPNPVPSPSPLTSFASCAEKGPVPEIQALQOEISISGCKMRLSYLSSTTPGVKSVLR 1077
1077 ISLTHTPTIPFNLMKVHLMVAVEGRFKWFAAAPDLSYFIWDTKTDVYNQKVFGLSEAFV 1136
1078 ISLTHTPTIPFNLMKVHLMVAVEGRFKWFAAAPDLSYFIWDTKTDVYNQKVFGLSEAFV 1137
1137 SVGYEYESCPDLILWEKETTTLQGYEIDASKLGWSLOKHALNTQS--GILHKGNGENQF 1195
1138 SVGYEYESCPDLILWEKETTTLQGYEIDASKLGWSLOKHALNTQS--GILHKGNGENQF 1197
1196 VSQQPPVTVGSMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGFNVRIRIF 1255
1198 VSQQPPVTVGSMNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGFNVRIRIF 1257
1256 PSGVNTNILEL--RNKDFRHSHPAHKYLLATDPMGSAVFLSDSNSRVRFKIKSTVVVKD 1313
1258 PSGVNTNILELRVNRKDFRHSHPAHKYLLATDPMGSAVFLSDSNSRVRFKIKSTVVVKD 1317
1314 LVKNEVVAGTGDQCLPFDDTRCGDGGKATLTPNR--GIVTDKGLYFVDDGTWIR 1370
1318 LVKNEVVAGTGDQCLPFDDTRCGDGGKATLTPNRGPGGIVTDKGLYFVDDGTWIR 1377
1371 RIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVR--LEWPTDLAINPMDNSLYVLDNN 1427
1378 RIDQNGIISTLLGSNDLTSARPLSCDSVMDISQVRQVHLEWPTDLAINPMDNSLYVLDNN 1437
1428 VVLQISENHQVRIVAGRPMHCOVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETD 1487
1438 VVLQISENHQVRIVAGRPMHCOVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETD 1497
1488 EKKINRIQVTTSGEISLVAGAPSCDCKDANCDGSGDDGYAKDAKLNTFSSLAVCAD 1547
1498 EKKINRIQVTTSGEISLVAGAPSCDCKDANCDGSGDDGYAKDAKLNTFSSLAVCAD 1557
1548 GELYVADLGNIRIRIRKXKPLNTQNMVELSSPIDQELYLFDTTGKHLYTQSILPTGYL 1607
1558 GELYVADLGNIRIRIRKXKPLNTQNMVELSSPIDQELYLFDTTGKHLYTQSILPTGYL 1617
1608 NYFTVTGDDITLITDNNGMNVNRDSTGMPLMLVVPDGOVYVMTGNSALKSVTTQG 1667
1618 NYFTVTGDDITLITDNNGMNVNRDSTGMPLMLVVPDGOVYVMTGNSALKSVTTQG 1677
1668 HELAMVTHGNSGLLATKSNENGMWTFTEYDPSFGLRTNVTFPTGQVSSFRSDTSSVHVQ 1727
1678 HELAMVTHGNSGLLATKSNENGMWTFTEYDPSFGLRTNVTFPTGQVSSFRSDTSSVHVQ 1737
1728 VETSSKDDVTITTNLSASGAFYTLLODVRSYVIGADGSLRLLLANGMEVALQTEPHLL 1787
1738 VETSSKDDVTITTNLSASGAFY----DQVRNSYVIGADGSLRLLLANGMEVALQTEPHLL 1793

```

1788 AGTVPVGRNVTLPIDNGLNLEWVRQKEQARGQVTVFGRSLR---VHNRNLSLDFD 1844
1794 AGTVPVGRNVTLPIDNGLNLEWVRQKEQARGQVTVFGRSLRVLQVHNRNLSLDFD 1853
1845 RVTRTEKIYDHRKFTLRILYDQAGPSLWSPSRRLNGVNVTVSPGGYIAGIQRGIMSR 1904
1854 RVTRTEKIYDHRKFTLRILYDQAGPSLWSPSRRLNGVNVTVSPGGYIAGIQRGIMSR 1913
1905 MEYDOAGSITSRIFADGKTWSYVLLKSMVLLHSQROQIFBDDKNDRLSSVTMNVASQ 1964
1914 MEYDOAGSITSRIFADGKTWSYVLLKAGV-----BFDKNDRLSSVTMNVASQ 1962
1965 TLETIRSVGYRNIIYQPEGNASVIOFTEDGHLHHTFVLGTGRRVIYKGLSKLAETL 2024
1963 TLETIRSVGYRNIIYQPEGNASVIOFTEDGHLHHTFVLGTGRRVIYKGLSKLAETL 2022
2025 YDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYQIGPLIDROIIFRTEEGMVARFDYN 2084
2023 YDTTKVSFTYDETAGMLKTINLQNEGFTCTIRYQIGPLIDROIIFRTEEGMVARFDYN 2082
2085 YDNSFRVTSQVAINETPLIDLYRVDVSGKTEQFGKGVIIYDINQIITAVMTHTKH 2144
2083 YDNSFRVTSQVAINETPLIDLYRVDVSGKTEQFGKGVIIYDINQIITAVMTHTKH 2142
2145 FDAYGRVKEVQYEIFRSLMTVMTVQVDMGRVVKKELKVGPNANTTRYSEYDADQLOT 2204
2143 FDAYGRVKEVQYEIFRSLMTVMTVQVDMGRVVKKELKVGPNANTTRYSEYDADQLOT 2202
2205 VSINDKPLWYSVDLGNLHLLSPGNSARLTPLRYDIRITRLGVQVQKMDGFLROR 2264
2203 VSINDKPLWYSVDLGNLHLLSPGNSARLTPLRYDIRITRLGVQVQKMDGFLROR 2262
2265 GGDIFEYNSAGLLIKAYNRAGSWSRYRVDGLGRVSSKSSHHLQFFYADLTNPTKVT 2324
2263 GGDIFEYNSAGLLIKAYNRAGSWSRYRVDGLGRVSSKSSHHLQFFYADLTNPTKVT 2322
2325 HLYNHSSEITSYLDYDQHLFAMELSSGDEFFIACDNIGTPLAVSGTGLMIKQILYTA 2384
2323 HLYNHSSEITSYLDYDQHLFAMELSSGDEFFIACDNIGTPLAVSGTGLMIKQILYTA 2382
2385 YGEIYMDTNPFIIGYHGLYDPLTKLVHMGRRDYDVLGRWTSPPDHLMKHLSSNV 2444
2383 YGEIYMDTNPFIIGYHGLYDPLTKLVHMGRRDYDVLGRWTSPPDHLMKHLSSNV 2442
2445 MPFNLYMFKNNPISNSQDIKCPMTDVNSWLLTFGQLHNVIPGYKPKDMDAMEPSYELI 2504
2443 MPFNLYMFKNNPISNSQDIKCPMTDVNSWLLTFGQLHNVIPGYKPKDMDAMEPSYELI 2502
2505 HTQMKTOEWNSK-----SILGVQCEVQKQKAFVTLERFDQLYGSTTSCQAPKT 2556
2503 HTQMKTOEWNSKVIIPAGQCSILGVQCEVQKQKAFVTLERFDQLYGSTTSCQAPKT 2562
2557 KKFASSGSVFGKGVKFKALDKGRVTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVT 2616
2563 KKFASSGSVFGKGVKFKALDKGRVTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVT 2622
2617 HYFVKPGSPGDLAILGLSGRKTLENGVNVTVSQINTVLNGTRRYTDIQLQYGALCN 2676
2623 HYFVKPGSPGDLAILGLSGRKTLENGVNVTVSQINTVLNGTRRYTDIQLQYGALCN 2682
2677 TRYGTTLDEKARVLELARQVAREQORLREGEGLRAWTGEXQOVLSTGRVQ 2736
2683 TRYGTTLDEKARVLELARQVAREQORLREGEGLRAWTGEXQOVLSTGRVQ 2742
2737 YDGFVVISVQYPELSDSANNIHFMOSENGRR 2769
2743 YDGFVVISVQYPELSDSANNIHFMOSENGRR 2775

```

```

; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Heriman, John I
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 82
; LENGTH: 2771
; TYPE: PRT
; ORGANISM: Mus musculus
; US-09-808-602-82

```

```

Query Match 97.6%; Score 14529; DB 9; Length 2771;
Best Local Similarity 97.0%; Pred. No. 0;
Matches 2688; Conservative 40; Mismatches 41; Indels 2; Gaps 1;

QY 1 MDVXERPYRLTTRRDAERRYTSSSADSBEGKAPQKSYSSSETLKAYDQADARLAYGSRV 60
DB 1 MDVXERPYRLTTRRDAERRYTSSSADSBEGKAPQKSYSSSETLKAYDQADARLAYGSRV 60
QY 61 KDIVPQAEFCRTGANFTLRELGLEVTTPHGLYKTDIGLPCGYSMGAGSDADMEAD 120
DB 61 KDMVPQAEFCRTGANFTLRELGLEVTTPHGLYKTDIGLPCGYSMGAGSDADMEAD 120
QY 121 TVLSPEHPVRLWGRSTSRSSCLSSRANSNLITDTEHENTETDHPGGLQNHARLTPP 180
DB 121 TVLSPEHPVRLWGRSTSRSSCLSSRANSNLITDTEHENTETDHPGGLQNHARLTPP 180
QY 181 PPLSHAHTPNQOHAASINSLNRGNFTPRSNPSPAPTDSLSGEPGAGQAPAHQENWL 240
DB 181 PPLSHAHTPNQOHAASINSLNRGNFTPRSNPSPAPTDSLSGEPGAGQAPAHQENWL 240
QY 241 LNSNIPLETNLGKQPFGLTQDNLIEMDILGASRHDGAYSDGHFLFKPGTSLFCTTS 300
DB 241 LNSNIPLETNLGKQPFGLTQDNLIEMDILGASRHDGAYSDGHFLFKPGTSLFCTTS 300
QY 301 PGYPLTSTVYSPPPRPLPRSTFPAPFNLKPKSKYCNWKAALSALVIGATLVILLAYF 360
DB 301 PGYPLTSTVYSPPPRPLPRSTFPAPFNLKPKSKYCNWKAALSALVIGATLVILLAYF 360
QY 361 VAMHLFGLNMLHQPMEG--QMYEITETASSWPVPTDVSLYPSGCTGLETDPDRKGKGTTE 418
DB 361 VAMHLFGLNMLHQPMEGQMYEITETASSWPVPTDVSLYPSGCTGLETDPDRKGKGAEE 420
QY 419 GKPSFPEDSFIDSGEIDVGRASOKIPGTFWRQVFDHPVHLNPNVSLGKAALVGI 478
DB 419 GKPSFPEDSFIDSGEIDVGRASOKIPGTFWRQVFDHPVHLNPNVSLGKAALVGI 480
QY 479 YRKGGLPESHTQPFVELLDGRRLLTQEARSLCTPQSRGTVPSPSHETGFIQYLDGSI 538
DB 479 YRKGGLPESHTQPFVELLDGRRLLTQEARSLCTPQSRGTVPSPSHETGFIQYLDGSI 540
QY 539 WLLAFYNDGKSEVVSFLTTTAAIESVDNCPNCGDCISGTCFCFLGPGDGRASCP 598
DB 539 WLLAFYNDGKSEVVSFLTTTAAIESVDNCPNCGDCISGTCFCFLGPGDGRASCP 600
QY 599 VILCSNGQYMKGRCLCHSGWKGAECVPTNQICIDVACSNHGTCITGTICNPGYKGSCE 658
DB 601 VILCSNGQYMKGRCLCHSGWKGAECVPTNQICIDVACSNHGTCITGTICNPGYKGSCE 660

```

QY	659	EVDGMDPTCSGRGVCVRGEBCHFCVGMGGTNCETPRATCLDQCQSGHGTFLPDTGLCSDPS	718
Db	661	EVDGMDPTCSGRGVCVRGEBCHFCVGMGGTNCETPRATCLDQCQSGHGTFLPDTGLCNCDS	720
QY	719	WTGHDGSIIECAADCGHGVCGVGTCTCEDGWMGAACDQACHPRCAEHGTCRDGKCECS	778
Db	721	WTGHDGSIIECAADCGHGVCGVGTCTCEDGWMGAACDQACHPRCAEHGTCRDGKCECS	780
QY	779	PGWNGBHCTIAHYLDRLVYKGGCPGLNGNGRCTLDLNGWCHVCQLGWRGAGCDTSMETAC	838
Db	781	PGWNGBHCTIAHYLDRLVYKGGCPGLNGNGRCTLDLNGWCHVCQLGWRGAGCDTSMETGC	840
QY	839	GDSDKNDGDLVDCMDPDCCLQPLCHINPLCLGSPNPLDIOHTQVYVQOQNLHSEFYDRI	898
Db	841	GDSDKNDGDLVDCMDPDCCLQPLCHINPLCLGSPNPLDIOHTQVYVQOQNLHSEFYDRI	900
QY	899	KFLVGRDSTHIIPEGNEPFDGHCACVIRGQVMTSDGTFPLGVNLSFVNPNLFGYVTSRQDG	958
Db	901	KFLVGRDSTHIIPEGNEPFDGHCACVIRGQVMTSDGTFPLGVNLSFVNPNLFGYVTSRQDG	960
QY	959	SFDLVNNGGSIILRPERAPFIQEHTLALPMDRPFVMTETIMRHEENIPSCDLSNFAR	1018
Db	961	SFDLVNNGGSIILRPERAPFIQEHTLALPMDRPFVMTETIMRHEENIPSCDLSNFAR	1020
QY	1019	PNFVSPSPLTSPASSCAEKGPVPEIQALQEBISISGCKMRLSYLSSRTPGVKSVLRIS	1078
Db	1021	PNFVSPSPLTSPASSCAEKGPVPEIQALQEBISISGCKMRLSYLSSRTPGVKSVLRIS	1080
QY	1079	LTHPTTIPNMLKVLVAVVEGRIFRWFAAAPDLSYFIFWKTVDVYVQVFLSEAFVSU	1138
Db	1081	LTHPTTIPNMLKVLVAVVEGRIFRWFAAAPDLSYFIFWKTVDVYVQVFLSEAFVSU	1140
QY	1139	GYEYESCPDLILWEKKTAVLQGEIDASKLGWSLQKHALNIQSGILHKGNGENQFVSQ	1198
Db	1141	GYEYESCPDLILWEKKTAVLQGEIDASKLGWSLQKHALNIQSGILHKGNGENQFVSQ	1200
QY	1199	QPPVIGSIMNGRRRISCPSCNGLADGNKLLAPVALTCGSDGLYVGFNFIIRIFPSG	1258
Db	1201	QPPVIGSIMNGRRRISCPSCNGLADGNKLLAPVALTCGSDGLYVGFNFIIRIFPSG	1260
QY	1259	NVTNILELRNKDPRHSHSPAHKYIYLATDPMGSAVFLSDNSRRVFKIYSTVVKOLVKN	1318
Db	1261	NVTNILELRNKDPRHSHSPAHKYIYLATDPMGSAVFLSDNSRRVFKIYSTVVKOLVKN	1320
QY	1319	EVVAGTGDQCLPDDTRCGDGGKATEATLTNPRGITVDXFGILYFVDGTMIRRIDONGII	1378
Db	1321	EVVAGTGDQCLPDDTRCGDGGKATEATLTNPRGITVDXFGILYFVDGTMIRRIDONGII	1380
QY	1379	STLLGNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYLDNNVLIQISENHQV	1438
Db	1381	STLLGNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYLDNNVLIQISENHQV	1440
QY	1439	RIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIQVY	1498
Db	1441	RIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIQVY	1500
QY	1499	TSGEISLVAGAPGCDCKNDANCDGSDGAGYAKDAKLNTSPSSLAYCADGELYVADLGN	1558
Db	1501	TSGEISLVAGAPGCDCKNDANCDGSDGAGYAKDAKLNTSPSSLAYCADGELYVADLGN	1560
QY	1559	RIFIRKPKPLNTQNYELSSPIDQELYLFDFTGKHLVTSQSLPTGDYLYNYTYTGDDGI	1618
Db	1561	RIFIRKPKPLNTQNYELSSPIDQELYLFDFTGKHLVTSQSLPTGDYLYNYTYTGDDGI	1620
QY	1619	TLITDNNGNMNVVRDSTGMPLMLVVPDGOVYVMTGMTNSALKSVTTQGHELAMMTYHGN	1678
Db	1621	TLITDNNGNMNVVRDSTGMPLMLVVPDGOVYVMTGMTNSALKSVTTQGHELAMMTYHGN	1680
QY	1679	SGLLATKSNENGWTTTFEYDPSFORLNTVPTQGVSSFRSDTSSVHVQVETSSKDDVTI	1738
Db	1681	SGLLATKSNENGWTTTFEYDPSFORLNTVPTQGVSSFRSDTSSVHVQVETSSKDDVTI	1740
QY	1739	TTNLSAGAFYTLQDQVRNSYIIGADGSLRULLANGMEVALQTEPHLLAGTVNPTVGKR	1798

RESULT 5

Db	1741	TTNLSAGAFYTLQDQVRNSYIIGADGSLRULLANGMEVALQTEPHLLAGTVNPTVGKR	1800
QY	1799	NVTLPIDNGLNLVWVRQKKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRK	1858
Db	1801	NVTLPIDNGLNLVWVRQKKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRK	1860
QY	1859	FTLRILYDQAGRPSLWSPSSRLNGVNTYSPGGVIAGIQRMERMEYDQAGRITSRIF	1918
Db	1861	FTLRILYDQAGRPSLWSPSSRLNGVNTYSPGGVIAGIQRMERMEYDQAGRITSRIF	1920
QY	1919	ADGKTWSYTYLEKSMVLLHSHQRYIFEFDKNDRLSVMTMENVARQTLTETRSVGYVNI	1978
Db	1921	ADGKTWSYTYLEKSMVLLHSHQRYIFEFDKNDRLSVMTMENVARQTLTETRSVGYVNI	1980
QY	1979	YQPEGNASVIOQDFTEDGHLHTFYLGTGRVVIYKYGKLSKLAETLYDTTKVSFTYETA	2038
Db	1981	YQPEGNASVIOQDFTEDGHLHTFYLGTGRVVIYKYGKLSKLAETLYDTTKVSFTYETA	2040
QY	2039	GMLKTNLONEGFCTTIRYQIIGPLIDROIFPRFTEEGVNARPDYNDNSFRVTSMAQVI	2098
Db	2041	GMLKTNLONEGFCTTIRYQIIGPLIDROIFPRFTEEGVNARPDYNDNSFRVTSMAQVI	2100
QY	2099	NETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITAVMTHTKHFDAYGRMKEVQVEI	2158
Db	2101	NETPLPIDLYRYDDVSGKTEQFGKFGVIYDINQIITAVMTHTKHFDAYGRMKEVQVEI	2160
QY	2159	FRSLMYMTVQYDNMGRVVKELKVGYPYANTTRYSYEDADGQLOTQVINDKPLWRSYD	2218
Db	2161	FRSLMYMTVQYDNMGRVVKELKVGYPYANTTRYSYEDADGQLOTQVINDKPLWRSYD	2220
QY	2219	LNGNLHLLSPGNSARLTPRLYDIDRITRLGDVOYKMDDEGFLRQRGDIPYNSAGLLI	2278
Db	2221	LNGNLHLLSPGNSARLTPRLYDIDRITRLGDVOYKMDDEGFLRQRGDIPYNSAGLLI	2280
QY	2279	KAYNRAGSWSVRYRYDGLGRVSSKSHHLPFYADLTNPTKVTHLYNHSSEITSLY	2338
Db	2281	KAYNRAGSWSVRYRYDGLGRVSSKSHHLPFYADLTNPTKVTHLYNHSSEITSLY	2340
QY	2339	YDLQGHLPAMELSGDEFIACDNIPTPLAVFSCTGLMIKQIILYAYGEIYMDTNPQOI	2398
Db	2341	YDLQGHLPAMELSGDEFIACDNIPTPLAVFSCTGLMIKQIILYAYGEIYMDTNPQOI	2400
QY	2399	IIGYHGGLYDPLTKLVHMGREDYDLVLAGRWTSPDHELMKHLSSSNVMPFNLVYMKNNPI	2458
Db	2401	IIGYHGGLYDPLTKLVHMGREDYDLVLAGRWTSPDHELMKHLSSSNVMPFNLVYMKNNPI	2460
QY	2459	SNSQDIKCFMTDVNSWLLTFQGLHNVIPGYPKPDMAMEPSYELIHTQMTQEWDSKS	2518
Db	2461	SNSQDIKCFMTDVNSWLLTFQGLHNVIPGYPKPDMAMEPSYELIHTQMTQEWDSKS	2520
QY	2519	ILGVQCEVOKQLKAFVTLERFDQLYGSTITSQCAPKTKFPASSGVFGKGVKALKDGR	2578
Db	2521	ILGVQCEVOKQLKAFVTLERFDQLYGSTITSQCAPKTKFPASSGVFGKGVKALKDGR	2580
QY	2579	VTTDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVDPHYVFKPSPSEGDIAIILGSGR	2638
Db	2581	VTTDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVDPHYVFKPSPSEGDIAIILGSGR	2640
QY	2639	RTLENGVNTVYSQINTVLNGSTRYTDIOLOYGALCNTRVGTLLDDEKARVLELQRA	2698
Db	2641	RTLENGVNTVYSQINTVLNGSTRYTDIOLOYGALCNTRVGTLLDDEKARVLELQRA	2700
QY	2699	VRQAWAREQORLRBEGEGELRAWTEGKQVVLSTGRVQYDGGFFVIVSEYQPELSDSANNI	2758
Db	2701	VRQAWAREQORLRBEGEGELRAWTEGKQVVLSTGRVQYDGGFFVIVSEYQPELSDSANNI	2760
QY	2759	HFMQSEMGRR	2769
Db	2761	HFMQSEMGRR	2771

US-09-800-198-70

; Sequence 70, Application US/09800198
; Publication No. US20030087816A1

; GENERAL INFORMATION:

; APPLICANT: Vernet, Cornie AM

; APPLICANT: Fernandes, Elma

; APPLICANT: Shimkets, Richard A

; APPLICANT: Herrmann, John L

; APPLICANT: Majumder, Kumud

; APPLICANT: Mishra, Vishna

; APPLICANT: Mezes, Peter S

; APPLICANT: Rastelli, Luca

; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

; FILE REFERENCE: 15966-697

; CURRENT APPLICATION NUMBER: US/09/800,198

; CURRENT FILING DATE: 2001-03-05

; PRIOR APPLICATION NUMBER: 60/186,596

; PRIOR FILING DATE: 2000-03-03

; NUMBER OF SEQ ID NOS: 98

; SOFTWARE: Patentin Ver. 2.1

; SEQ ID NO 70

; LENGTH: 2771

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-800-198-70

Query Match 97.6%; Score 14529; DB 10; Length 2771;

Best Local Similarity 97.0%; Pred. No. 0;

Matches 2688; Conservative 40; Mismatches 41; Indels 2; Gaps 1;

Qy 1 MDVKERKPYSLRRRDAERRYSSADSEEGKAPOKSYSSSETLKAYDQDARLAYGSRV 60

Db 1 MDVKERKPYSLRRRDAERRYSSADSEEGKAPOKSYSSSETLKAYDQDARLAYGSRV 60

Qy 61 KDIVPOAEFFCTGANFTLRLEGLBEVTPPHGTLYRTDGLPQCGYSGAGSDADMEAD 120

Db 61 KDVVPOAEFFCTGTNFTLRLEGLGEMTPPHGTLYRTDGLPQCGYSGAGSDADLEAD 120

Qy 121 TVLSPEHPVRLWGRSTRGRSSCLSRANSNLTLTDTEHENTETDHPGLQNHARLRTPP 180

Db 121 TVLSPEHPVRLWGRSTRGRSSCLSRANSNLTLTDTEHENTETDHPSSLQNHPRLRTPP 180

Qy 181 PPLSHAFTPNQHAASINSINLRGNFTPRSNPAPPTDHSLSGEPAGGAQOEPAHAQENWL 240

Db 181 PPLPHAFTPNQHAASINSINLRGNFTPRSNPAPPTDHSLSGEPAGGAPSAQEPHAQDNWL 240

Qy 241 LNSNIPLERNLKGQFLGTLDQNLLEMDILGASRDHGYSDGHFLFKPGGTSPLFCCTS 300

Db 241 LNSNIPLERNLKGQFLGTLDQNLLEMDILGASRDHGYSDGHFLFKPGGTSPLFCCTS 300

Qy 301 PGYPLTSSVTVSPPPRLPRSTFARPAFLNKKPSKYCNWKAALSAIVISATLVILLAYF 360

Db 301 PGYPLTSSVTVSPPPRLPRSTFARPAFLNKKPSKYCNWKAALSAILLISATLVILLAYF 360

Qy 361 VAMHLFGLNHLQPMEG--QMYEITDSTASSVPVPTDVSILYSGGTGLTTPDRKKGYTE 418

Db 361 VAMHLFGLNHLQPMEGQMYEITDSTASSVPVPTDVSILYSGGTGLTTPDRKKGAAE 420

Qy 419 GKPSPPFPDSFIDSIEIDVGRASOKIPPGTFWRQSVFIDHPVHLKFNVSILGKAALVGI 478

Db 421 GKPSLFPDSFIDSIEIDVGRASOKIPPGTFWRQSVFIDHPVHLKFNVSILGKAALVGI 480

Qy 479 YGRKGLPPSHQTQDFVELLDGRLLTQEARSLLEGTPRQSRGTVPPSSSHETGFIQYLDGI 538

Db 481 YGRKGLPPSHQTQDFVELLDGRLLTQEARSLLEGTPRQSRGTVPPSSSHETGFIQYLDGI 540

Qy 539 WHLAFYNDGKESVWFLTTATIESVNCNFCNVCNGDCISGTCFPLGPDCCGRASCP 598

Db 541 WHLAFYNDGKESVWFLTTATIESVNCNFCNVCNVCNGDCISGTCFPLGPDCCGRASCP 600

Qy 599 VLGSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCITGTICNPYKGESCE 658

Db 601 VLGSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSNHGTCIMGTICINPYKGESCE 660

Qy 659 EVDGMDPTCSGRGVCVVRGECHECFVVGWGTNCETPRATCLDQCSCGHGTFLPDTGLCSCDPS 718

Db 661 EVDGMDPTCSSRGVCVVRGECHECFVVGWGTNCETPRATCLDQCSCGHGTFLPDTGLCNCDCPS 720

Qy 719 WTGHDCSIEICAADCGGHGVCVVGTCRCEDGWGAACDQACPRCAEHGTCRDGKCECS 778

Db 721 WTGHDCSIEICAADCGGHGVCVVGTCRCEDGWGAACDQACPRCAEHGTCRDGKCECS 780

Qy 779 PGWNGEHCITIAHYLDREVYKGCPCGLCNGNCRCTLDLNGMHCVCQLGMRGAGCDDTSMETAC 838

Db 781 PGWNGEHCITIAHYLDREVYKGCPCGLCNGNCRCTLDLNGMHCVCQLGMRGAGCDDTSMETGC 840

Qy 839 GDSKNDGDGLVDCMDPDCCLQPLCHINPLCLSPNPLDIIQETQVPSQNLHSPYDRI 898

Db 841 GDGKNDGDGLVDCMDPDCCLQPLCHINPLCLSPNPLDIIQETQVPSQNLHSPYDRI 900

Qy 899 KFLVGRDSTHIIIGENPFDCGHACVIRGOVMTSDGTPLVGVNISFVNNPLFGYTIISRDQG 958

Db 901 KFLVGRDSTHIIIGENPFDCGHACVIRGOVMTSDGTPLVGVNISFVNNPLFGYTIISRDQG 960

Qy 959 SFDLVTVNGGISIIILRPERAPFITQEHITLWLPWRPFVMTIIMRHEENIIPSCDLSNFAR 1018

Db 961 SFDLVTVNGGISIIILRPERAPFITQEHITLWLPWRPFVMTIIMRHEENIIPSCDLSNFAR 1020

Qy 1019 PNPVVSPLTSPASSCAEKGPIVPEIQALQOEIISISGCKWRLSYLSSTPGYKSVLRIS 1078

Db 1021 PNPVVSPLTSPASSCAEKGPIVPEIQALQOEIIVIAAGCKWRLSYLSSTPGYKSVLRIS 1080

Qy 1079 LTHPTIPFNLKVMHLMVAVVEGRFLFRKWFAPAAAPDLSYFFIMDKTDVYNQKVFGESEAFVSU 1138

Db 1081 LTHPTIPFNLKVMHLMVAVVEGRFLFRKWFAPAAAPDLSYFFIMDKTDVYNQKVFGESEAFVSU 1140

Qy 1139 GYEYESCFDILILEKRTTVLQGYEIDASKLGGHSLDKHHAALNIQSGILHKGNGENOFVSQ 1198

Db 1141 GYEYESCFDILILEKRTAVLQGYEIDASKLGGHSLDKHHAALNIQSGILHKGNGENOFVSQ 1200

Qy 1199 QPPVIGSIMGNGRRRSISCPSCNGGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFFPSG 1258

Db 1201 QPPVIGSIMGNGRRRSISCPSCNGGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFFPSG 1260

Qy 1259 NVTNILELRNKDPRHSHSPAHKYLLATDPMGAVFLSDNSRRVFKIKSTVVVKDLVKNS 1318

Db 1261 NVTNILEMRNKDPRHSHSPAHKYLLATDPMGAVFLSDNSRRVFKIKSTVVVKDLVKNS 1320

Qy 1319 EVVAGTGDQCLPFDDTRCGDGKATEATLTPRGITVDKFLIYFVDGTMIRRIDQNGII 1378

Db 1321 EVVAGTGDQCLPFDDTRCGDGKATEATLTPRGITVDKFLIYFVDGTMIRRIDQNGII 1380

Qy 1379 STLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNVLOISENHQV 1438

Db 1381 STLGSNDLTSARPLSCDSVMEISQVRLEWPTDLAINPMDNSLYVLDNNVLOISENHQV 1440

Qy 1439 RIVAGRPMHCOVPGIDHFLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIRQVT 1498

Db 1441 RIVAGRPMHCOVPGIDHFLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIRQVT 1500

Qy 1499 TSGEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLANTPSSLAYCAGDELYVADLNI 1558

Db 1501 TSGEISLVAGAPSGCDCKNDANCDGSGDDGYAKDAKLANTPSSLAYCAGDELYVADLNI 1560

Qy 1559 RIRIRKXNKPLNTQNNYELSSPIDQELYLFDITGHELYTQSLPTGDIYLYNFYTTGGDI 1618

Db 1561 RIRIRKXNKPLNTQNNYELSSPIDQELYLFDITGHELYTQSLPTGDIYLYNFYTTGGDI 1620

Qy 1619 TLITDNNGNMNVRRRDSGTGMPLMLVVPDGOVYVMTGNTSALKSVTTQGHFLAMMTYHGN 1678

Db 1621 THITDNNGNMNVRRRDSGTGMPLMLVVPDGOVYVMTGNTSALKSVTTQGHFLAMMTYHGN 1680

Qy 1679 SGLLATXSNENGWTTTFEYDSFGRLNTVPTTQVSSFRSDTDSVHVQVETSSKDDVTI 1738

Db 1681 SGLLATXSNENGWTTTFEYDSFGRLNTVPTTQVSSFRSDTDSVHVQVETSSKDDVTI 1740

QY 1739 TTNLASGAFYTLQOVRSYIIGADGSLRLLLANGMEVALQTEPHLLAGTVPVGR 1798
 Db 1741 TTNLASGAFYTLQOVRSYIIGADGSLRLLLANGMEVALQTEPHLLAGTVPVGR 1800
 QY 1799 NVTLPIDNGLNVWRQKEQARQGVTVFGRRLRVHNRNLLSDFRVRTTEKIYDDHRK 1858
 Db 1801 NVTLPIDNGLNVWRQKEQARQGVTVFGRRLRVHNRNLLSDFRVRTTEKIYDDHRK 1860
 QY 1859 FTIRILYDQAGRLSPSSRLGNVNTVSPGVYIAGIQRGINSEMEYDQAGRTSRIF 1918
 Db 1861 FTIRILYDQAGRLSPSSRLGNVNTVSPGVYIAGIQRGINSEMEYDQAGRTSRIF 1920
 QY 1919 ADGKTSYTYLEKSMVLLHSQRYIFEDKNDRLSSVTMPNVARQTLTIRSVGYRNI 1978
 Db 1921 ADGKTSYTYLEKSMVLLHSQRYIFEDKNDRLSSVTMPNVARQTLTIRSVGYRNI 1980
 QY 1979 YQPEGNASVIQDTEGHLHFTYGTGRVYKYKLSKLAETLYDTTKVSYFYDETA 2038
 Db 1981 YQPEGNASVIQDTEGHLHFTYGTGRVYKYKLSKLAETLYDTTKVSYFYDETA 2040
 QY 2039 GMLKTINLQNEGTCTIRYRQIGPLIDRQIFRTEGWNARPDYNDNSFRVTSQCAVI 2098
 Db 2041 GMLKTINLQNEGTCTIRYRQIGPLIDRQIFRTEGWNARPDYNDNSFRVTSQCAVI 2100
 QY 2099 NETPLPIDLYRVDVSGKTEQFGKFGVIYYDINQIITTAVMTHTKHFDAYGRMKEVQYBI 2158
 Db 2101 NETPLPIDLYRVDVSGKTEQFGKFGVIYYDINQIITTAVMTHTKHFDAYGRMKEVQYBI 2160
 QY 2159 FRSLMYMTVOYDNMGVVKKEIKVGPYANTYRYSVEYDADGLOQTVSINDKPLWYSYD 2218
 Db 2161 FRSLMYMTVOYDNMGVVKKEIKVGPYANTYRYSVEYDADGLOQTVSINDKPLWYSYD 2220
 QY 2219 LKGNLHLLSPGNARLTPLDYRIDRITRLGVDQYKMDGDFLRQGRGDI FEYNSAGLLI 2278
 Db 2221 LKGNLHLLSPGNARLTPLDYRIDRITRLGVDQYKMDGDFLRQGRGDI FEYNSAGLLI 2280
 QY 2279 KAYNRAGSMVRYRVDGLGRVSKSHSHLOFFFYADLTNPTKVTHLYNHSSEITSY 2338
 Db 2281 KAYNRAGSMVRYRVDGLGRVSKSHSHLOFFFYADLTNPTKVTHLYNHSSEITSY 2340
 QY 2339 YDIQGLHFMELSSGDFYIACDNIGTFLAVFSGTGLMIKQILYIYAGIYMDTNFNFOI 2398
 Db 2341 YDIQGLHFMELSSGDFYIACDNIGTFLAVFSGTGLMIKQILYIYAGIYMDTNFNFOI 2400
 QY 2399 IIGYHGLYDPLTKLVHMGRRDYLVLGRWTSYDHELWKLSSNVMPNLYMFKNNPI 2458
 Db 2401 IIGYHGLYDPLTKLVHMGRRDYLVLGRWTSYDHELWKLSSNVMPNLYMFKNNPI 2460
 QY 2459 SNSODIKCFMTDVNSWLLTFGLHNVIPGYPKPDMAMPSEYELIHTQMTQEWDSKS 2518
 Db 2461 SNSODIKCFMTDVNSWLLTFGLHNVIPGYPKPDMAMPSEYELIHTQMTQEWDSKS 2520
 QY 2519 ILGVQCEVQKQKAFVTLERFDQLYGTTITSCQAPKTKFASGSGVFGKVKFALKDGR 2578
 Db 2521 ILGVQCEVQKQKAFVTLERFDQLYGTTITSCQAPKTKFASGSGVFGKVKFALKDGR 2580
 QY 2579 VTTDIISVANEDGERVAAILNHAHYLENLHFTIDGVDTHTYFVKPGSEGLATILGSGGR 2638
 Db 2581 VTTDIISVANEDGERVAAILNHAHYLENLHFTIDGVDTHTYFVKPGSEGLATILGSGGR 2640
 QY 2639 RTLENGVNVTVSQINTVNGRTRYTDIQLQYGCALCINTYGTTLDEEKAARVLELARQRA 2698
 Db 2641 RTLENGVNVTVSQINTVNGRTRYTDIQLQYGCALCINTYGTTLDEEKAARVLELARQRA 2700
 QY 2699 VRQWAREOQRLREGEGLRAWTGEGKQOVLSTGRVQGYDGFVIVSQYPELSDSANNI 2758
 Db 2701 VRQWAREOQRLREGEGLRAWTGEGKQOVLSTGRVQGYDGFVIVSQYPELSDSANNI 2760
 QY 2759 HFMRQSEMGR 2769
 Db 2761 HFMRQSEMGR 2771

RESULT 6
 US-10-042-865-2
 ; Sequence 2, Application US/10042865
 ; Publication No. US2004029216A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Padigaru, Muralidhara
 ; APPLICANT: Li, Li
 ; APPLICANT: Zerhusen, Bryan D
 ; APPLICANT: Casman, Stacie J
 ; APPLICANT: Shenoy, Suresh G
 ; APPLICANT: Spytek, Kimberly
 ; APPLICANT: Zhong, Mei
 ; APPLICANT: Gangolli, Esha A
 ; APPLICANT: Burgess, Catherine E
 ; APPLICANT: Patturajan, Meera
 ; APPLICANT: Vernet, Corine A.M
 ; APPLICANT: Taylor, Sarah
 ; APPLICANT: Tchernev, Velizar T
 ; APPLICANT: Miller, Charles E
 ; APPLICANT: Guo, Xiaojia
 ; APPLICANT: Boldog, Ference L
 ; APPLICANT: Grosse, William M
 ; APPLICANT: Alsobrook II, John P
 ; APPLICANT: Gerlach, Valerie L
 ; APPLICANT: Edinger, Shlomit R
 ; APPLICANT: Rotherberg, Mark E
 ; APPLICANT: Ellerman, Karen
 ; APPLICANT: MacDougall, John
 ; APPLICANT: Malyankar, Uriel M
 ; APPLICANT: Millet, Isabelle
 ; APPLICANT: Peyman, John
 ; APPLICANT: Smithson, Glennda
 ; APPLICANT: Gunther, Erik
 ; APPLICANT: Stone, David
 ; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
 ; FILE REFERENCE: 21402-537
 ; CURRENT APPLICATION NUMBER: US/10/042,865
 ; CURRENT FILING DATE: 2002-05-17
 ; PRIOR APPLICATION NUMBER: 60/260,417
 ; PRIOR FILING DATE: 2001-01-09
 ; PRIOR APPLICATION NUMBER: 60/260,831
 ; PRIOR FILING DATE: 2001-01-10
 ; PRIOR APPLICATION NUMBER: 60/272,338
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR APPLICATION NUMBER: 60/274,876
 ; PRIOR FILING DATE: 2001-03-09
 ; PRIOR APPLICATION NUMBER: 60/284,704
 ; PRIOR FILING DATE: 2001-04-18
 ; NUMBER OF SEQ ID NOS: 264
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 2794
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-042-865-2
 ; Query Match 96.8%; Score 14406.5; DB 12; Length 2794;
 ; Best Local Similarity 96.7%; Pred. No. 0;
 ; Matches 2704; Conservative 21; Mismatches 41; Indels 31; Gaps 11;
 QY 1 MDVKERKPYRSLTRRDAERYTSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60
 Db 1 MDVKERKPYRSLTRRDAERYTSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV 60
 QY 61 KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLRYRTDGLPQCGYMGAGSDADMEAD 120
 Db 61 KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLRYRTDGLPQCGYMGAGSDADMEAD 120
 QY 121 TVLSPEHPVRLWGRSTRGRSSCLSSRANSNLTLTDEHENTETDHPGCLQNHARLRTPP 180
 Db 121 TVLSPEHPVRLWGRSTRGRSSCLSSRANSNLTLTDEHENTETDHPGCLQNHARLRTPP 180

QY	181	PPLSHAHTPNCHRAASINLNNGNFTPRSNPSPAPTDHSLSGBPAGGAEPAHAQENWL	240
DB	181	PPLSHAHTPNQHAASINLNNGNFTPRSNPSPAPTDHSLSGBPAGGAEPAHAQENWL	240
QY	241	LNSNIPLETNLGKQPFGLTQDNLEMDILGASRDHGDAYSQDGHFLPKPGGTSPLFCCTS	300
DB	241	LNSNIPLETNLGKQPFGLTQDNLEMDILGASRDHGDAYSQDGHFLPKPGGTSPLFCCTS	300
QY	301	PGYPLTSTSYSPRPPLRSTTARAFNLKPKSKYCNMKCAALSALVISATLIVLLAYP	360
DB	301	PGYPLTSTSYSPRPPLRSTTARAFNLKPKSKYCNMKCAALSALVISATLIVLLAYP	360
QY	361	VAMELFGNLHQLPMQOMYEITETDASSWPVPTDVSPLYSGGTGLTDRKKGKTEGK	420
DB	361	VAMELFGNLHQLPMQOMYEITETDASSWPVPTDVSPLYSGGTGLTDRKKGKTEGK	420
QY	421	PSFPPEPDSIDSGEIDVGRASOKIPPGTFWRSQVFIIDHPVHLKENVSLGRAALVGIY	480
DB	421	PSFPPEPDSIDSGEIDVGRASOKIPPGTFWRSQVFIIDHPVHLKENVSLGRAALVGIY	480
QY	481	RKGLPPSHTQDFVELLDGRLLTQEARSLGTPROSRTGTPPSSSHETGFIQVLDGSIWH	540
DB	481	RKGLPPSHTQDFVELLDGRLLTQEARSLGTPROSRTGTPPSSSHETGFIQVLDGSIWH	540
QY	541	LAFYNDGKESVSVFLTTAIESVDNCPNCGYNGDCISGTCHCFGLGPDGGRASCPVL	600
DB	541	LAFYNDGKESVSVFLTTAIESVDNCPNCGYNGDCISGTCHCFGLGPDGGRASCPVL	597
QY	601	CSNGQYMKGRCLCHSGWKAECDDVPTNOCIDVACSNHGTCTGTGTCINPGYKGESCEB	660
DB	598	CSNGQYMKGRCLCHSGWKAECDDVPTNOCIDVACSNHGTCTGTGTCINPGYKGESCEB	657
QY	661	DCMDPTCSGVCVGRGCHCFVWGNGTNCETPRATCLDQCSHGTFPLDTGLCSDPSWT	720
DB	658	DCMDPTCSGVCVGRGCHCFVWGNGTNCETPRATCLDQCSHGTFPLDTGLCSDPSWT	717
QY	721	GHDCEIEICADCGGHGCVGTCRCEBDMWGAACDORACHPRCAEHGTCRCKGCEB	780
DB	718	GHDCEIEICADCGGHGCVGTCRCEBDMWGAACDORACHPRCAEHGTCRCKGCEB	777
QY	781	WNGEHCITII--AHYLDRLVVK--EGCPGLCNGNGRCTLDLNGWHVCVQLGWRGAGCDSMET	836
DB	778	WNGEHCITII--AHYLDRLVVK--EGCPGLCNGNGRCTLDLNGWHVCVQLGWRGAGCDSMET	837
QY	837	ACGSKNDGDLVDCMDPCCLOPLCHINPLCLGSPNPLDIIQETQVPSVQONLHSFYD	896
DB	838	ACGSKNDGDLVDCMDPCCLOPLCHINPLCLGSPNPLDIIQETQVPSVQONLHSFYD	897
QY	897	RIRKPLVGRDSTHIIIPGENPDPGGHACVIRQVMTSDGTPLVGNISFVANNPLFGYTISRQ	956
DB	898	RIRKPLVGRDSTHIIIPGENPDPGGHACVIRQVMTSDGTPLVGNISFVANNPLFGYTISRQ	957
QY	957	DGSPDLVNGGSIILIRFERAPITIOEHTLWLPWDRFFVWETIIMHEBNEIPSCDLSNF	1016
DB	958	DGSPDLVNGGSIILIRFERAPITIOEHTLWLPWDRFFVWETIIMHEBNEIPSCDLSNF	1017
QY	1017	ARNPVPVSPPLTSFASSCAEKGPVPEIQALQEEISISCKMRLSYLSRTPGYKSULR	1076
DB	1018	ARNPVPVSPPLTSFASSCAEKGPVPEIQALQEEISISCKMRLSYLSRTPGYKSULR	1077
QY	1077	ISLTHPTIPNLMKVHLWMAVEGRLEFRKWFAPAAAPDLSYFIMDKTDVYNQKVFGESEAFV	1136
DB	1078	ISLTHPTIPNLMKVHLWMAVEGRLEFRKWFAPAAAPDLSYFIMDKTDVYNQKVFGESEAFV	1137
QY	1137	SVGYEYESCDLILWEKRTVLQGYEIDASKGWSLDKHALNIOG--GILHKGGENOF	1195
DB	1138	SVGYEYESCDLILWEKRTVLQGYEIDASKGWSLDKHALNIOG--GILHKGGENOF	1197
QY	1196	VSQOPPVIGSIMGNGRRRSISCFSCNGLADGNKLLAPVALTCGSDGSLVYGDFNYIRRF	1255
DB	1198	VSQOPPVIGSIMGNGRRRSISCFSCNGLADGNKLLAPVALTCGSDGSLVYGDFNYIRRF	1257
QY	1256	PSGNVTNILEL--RNKDFRSHSHPAHKYIYLATDPMGAVFLSDNSNRVFKIKSTVVVKD	1313

DB	1258	PSGNVTNILELVRNKKDFRSHSHPAHKYIYLATDPMGAVFLSDNSNRVFKIKSTVVVKD	1317
QY	1314	LVKNSVAVAGTGDCLPDDTRCGDGKGAATEATLTPR--GITVDKFGGLIYFVDDGTWIR	1370
DB	1318	LVKNSVAVAGTGDCLPDDTRCGDGKGAATEATLTPRPGPGITVDKFGGLIYFVDDGTWIR	1377
QY	1371	RIDQNGIISTLGLSNDLTSARPLSCDSVMDISQVR--LEWPTDLAINPMDNSLYVLDDN	1427
DB	1378	RIDQNGIISTLGLSNDLTSARPLSCDSVMDISQVRQVHLEWPTDLAINPMDNSLYVLDDN	1437
QY	1428	VVLQISENHQVRIIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETD	1487
DB	1438	VVLQISENHQVRIIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETD	1497
QY	1488	EKKINRIRQVTTSGEISLVAGAPSGCDKNDANDCFSGDDGYAKDAKLANTPSSLAVCAD	1547
DB	1498	EKKINRIRQVTTSGEISLVAGAPSGCDKNDANDCFSGDDGYAKDAKLANTPSSLAVCAD	1557
QY	1548	GELYVADLGNIRIRFIRKNKPFLLNTQNYELSSPIDQELYLFDITGKHYTQSLPTGDYL	1607
DB	1558	GELYVADLGNIRIRFIRKNKPFLLNTQNYELSSPIDQELYLFDITGKHYTQSLPTGDYL	1617
QY	1608	YNFTYTGDDITLITDNNGNMNVRRDSTGMPLWLVVDPGVVWMTGNTSALKSVTTQG	1667
DB	1618	YNFTYTGDDITLITDNNGNMNVRRDSTGMPLWLVVDPGVVWMTGNTSALKSVTTQG	1677
QY	1668	HELAMTYVHGNSGLLATKSNENGWTTTFEYDVSFGRLLNVTFPTGQVSSFRSDTSSVHVQ	1727
DB	1678	HELAMTYVHGNSGLLATKSNENGWTTTFEYDVSFGRLLNVTFPTGQVSSFRSDTSSVHVQ	1737
QY	1728	VETSSKDDVTITNLASGAPYLLQDQVRNSYIIGADGSLRLLLLANGMEVALQTEPHLL	1787
DB	1738	VETSSKDDVTITNLASGAPYLLQDQVRNSYIIGADGSLRLLLLANGMEVALQTEPHLL	1797
QY	1788	AGTVNPTVGRKNVTLPTDNGNLNVEMRQKEARGQVTFVGRRLR--VHNRLLSLDLFD	1844
DB	1798	AGTVNPTVGRKNVTLPTDNGNLNVEMRQKEARGQVTFVGRRLRVLQVHNRLLSLDLFD	1857
QY	1845	RVTETEKIYDDHRTTLRLLYDQAGRPSLWSPSSRLNGVNVTVSPGYIAGIQRGIMSER	1904
DB	1858	RVTETEKIYDDHRTTLRLLYDQAGRPSLWSPSSRLNGVNVTVSPGYIAGIQRGIMSER	1917
QY	1905	MEYDQAGRIISRIIFADGKTSYTYLEK----SMVLLHSORQYIFEDKNDRLSSVTMPN	1960
DB	1918	MEYDQAGRIISRIIFADGKTSYTYLEKAGVQSWLLHSORQYIFEDKNDRLSSVTMPN	1977
QY	1961	VARQTLTIRSVGYRNIYQFPEGNASVIQDFTEDGHLHHTFVLGTGRRVYIKYKLSKL	2020
DB	1978	VARQTLTIRSVGYRNIYQFPEGNASVIQDFTEDGHLHHTFVLGTGRRVYIKYKLSKL	2037
QY	2021	AETLYDTTKVSFTYDETAGMLKTINLQNEGTCTIRYQIPLIDROIIFRTEEGMNAR	2080
DB	2038	AETLYDTTKVSFTYDETAGMLKTINLQNEGTCTIRYQIPLIDROIIFRTEEGMNAR	2097
QY	2081	FDYNDNSFRVTSQAVINETPLIDLYRYDDVSGKTEQFQKFGVIYDINQIITAVMT	2140
DB	2098	FDYNDNSFRVTSQAVINETPLIDLYRYDDVSGKTEQFQKFGVIYDINQIITAVMT	2157
QY	2141	HTKHFDAYGRMKEVQYBIFRSLMYWMTVQYDNNMGRVVKELKVGYPYANTTRYEYDADG	2200
DB	2158	HTKHFDAYGRMKEVQYBIFRSLMYWMTVQYDNNMGRVVKELKVGYPYANTTRYEYDADG	2217
QY	2201	QLQTVSINDPLWRYSDYDLNGLNHLSPGNSARLTPLRYDIRDRIIRLGDVQVMBEDGF	2260
DB	2218	QLQTVSINDPLWRYSDYDLNGLNHLSPGNSARLTPLRYDIRDRIIRLGDVQVMBEDGF	2277
QY	2261	LRQGGDIIFEYNSAGLLIKAYNRAGSVRYRYDGLGRRVSSKSSHHLHQFFYADLTNP	2320
DB	2278	LRQGGDIIFEYNSAGLLIKAYNRAGSVRYRYDGLGRRVSSKSSHHLHQFFYADLTNP	2337
QY	2321	TKVTHLYNHSSSETLSLYDLQGHLPAMELSSGDGEFYACDNIGTGPLAVFSGTGLMKOI	2380


```

Db 2338 TKVTHLYNHSSSITSLYYDLQHLFAMELSSGDEFFYIACDNIGTPLAVFSGTGLMIKI 2397
QY 2381 LYTAYGEIYNDINPNFQIIIGYHGGYDPLTKLVHMGRRDYDLVLAGRWTSPDHELMKHL 2440
Db 2398 LYTAYGEIYNDINPNFQIIIGYHGGYDPLTKLVHMGRRDYDLVLAGRWTSPDHELMKHL 2457
QY 2441 SNNMFPNLYMFKNWNPISNQDIKCFMTDVNSWLLTFQFQLHNVIPGYPKPDMDAMEPS 2500
Db 2458 SNNMFPNLYMFKNWNPISNQDIKCFMTDVNSWLLTFQFQLHNVIPGYPKPDMDAMEPS 2517
QY 2501 YELIHTQMKTOEWDNSK-----SILGVQCVRQKQKAFVTLERFDOLYSTITSQO 2552
Db 2518 YELIHTQMKTOEWDNSKVTPAQCCQILGVQCEVQKQKAFVTLERFDOLYSTITSQO 2577
QY 2553 APTKXKFASGVSFGKGVKFPALKDGRVTTDIISVANEDGRRVAAIINHAYLENLHFTTD 2612
Db 2578 APTKXKFASGVSFGKGVKFPALKDGRVTTDIISVANEDGRRVAAIINHAYLENLHFTTD 2637
QY 2613 GVDTHTVFKPGSEGDLAILGLSGGRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGA 2672
Db 2638 GVDTHTVFKPGSEGDLAILGLSGGRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGA 2697
QY 2673 LCLNTRYGTTLDDEKARVLELQRAVRQAWAREQOQLREGEGLRAWTGEXKQQVLSTG 2732
Db 2698 LCLNTRYGTTLDDEKARVLELQRAVRQAWAREQOQLREGEGLRAWTGEXKQQVLSTG 2757
QY 2733 RVQGYDGFVIVSQVPELSDSANNTHFMROSEMGR 2769
Db 2758 RVQGYDGFVIVSQVPELSDSANNTHFMROSEMGR 2794

```

RESULT 7

```

US-10-042-865-52
; Sequence 52, Application US/10042865
; Publication No. US20040029216A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan D
; APPLICANT: Casman, Stacie J
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zhong, Mei
; APPLICANT: Gangolli, Baha A
; APPLICANT: Burgess, Catherine E
; APPLICANT: Patturajan, Meera
; APPLICANT: Vernet, Corine A.M
; APPLICANT: Taylor, Sarah
; APPLICANT: Tchernev, Velizar T
; APPLICANT: Miller, Charles E
; APPLICANT: Guo, Xiaojia
; APPLICANT: Boldog, Ference L
; APPLICANT: Gross, William M
; APPLICANT: Alsbrook II, John P
; APPLICANT: Gerlach, Valerie L
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Rothenberg, Mark E
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John
; APPLICANT: Smithson, Glenda
; APPLICANT: Gunther, Erik
; APPLICANT: Stone, David
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of
; FILE REFERENCE: 21402-537
; CURRENT APPLICATION NUMBER: US/10/042.865
; CURRENT FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: 60/260,417
; PRIOR FILING DATE: 2001-01-09
; PRIOR APPLICATION NUMBER: 60/260,831

```

```

; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 60/272,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/274,876
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/284,704
; PRIOR FILING DATE: 2001-04-18
; NUMBER OF SEQ ID NOS: 264
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 52
; LENGTH: 2715
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-042-865-52

Query Match 68.7%; Score 10227; DB 12; Length 2715;
Best Local Similarity 66.7%; Pred. No. 0;
Matches 1864; Conservative 370; Mismatches 457; Indels 102; Gaps 19;

QY 1 MDVKEKPYRSLTR-RRDABRRYTSSSADSEEGKAP-QKSYSSSETLKAYQDQ-ARLAYG 57
Db 1 MDVKERRYCSLTKSRREKERRYINSSADNEECRVPTQKSYSSSETLKAFDHDYSRLLYG 60
QY 58 SRVKDIVEPQEAEEFRTGANTFLRELGLEEVTPPHGTLRYRFDIGLPQCGYMGAGSDADM 117
Db 61 NRVKDLVHREADEVTRQGNFTLQLVGYESATRRGVAFCAEMGLPHRGYSISAGSDADT 120
QY 118 EADTVLSPEHPVRLWGRSTRSGRSCJSSRANSNLTLTDTEHEN---TETHPGGLQNH 174
Db 121 ENEAVMSPEHAMRLWGRGVKSGRSCJSSRANSNLTLTDTEHENRSESESQPPNNPCQP 180
QY 175 RLRTPPPLSHAHTPNQHHAAASINSLNRGNFTPRSNPSAPTDHSLSGEPAGGAQ---- 230
Db 181 TLQPLPP--SHKQHPAQHH-PSITSLNRNSLTNRNQSFAP-----PAALPAELQTT 229
QY 231 -EPHAQENWLLNSNIPLETNGLKQPPLGLTLDNLIEMDILGASRHDCAYSDDGHFLEKP 289
Db 230 PESVQLQDSWVLGNSVPLESR-----HFLPKT 256
QY 290 G-GTSPFLCTTSPGYPLTSSTVYSPPPRPLPRSTFARPAFNLKPKSKYCNKCAALSAIV 348
Db 257 GTGTTPLFSTATPGYTMASGVSYPPTPLPRNLTLSRASFKEFKSKYCNKCAALCAVG 316
QY 349 ISATLIVLLAFVAMHLFGLNHLQPMQWYME---ITEDTASSVPVTDVSLVPSGGTG 405
Db 317 VSVLLAILLSYFIAMHLFGLNHLQPMQWYME---ITEDTASSVPVTDVSLVPSGGTG 368
QY 406 LETPDKGKGITTEGKPSFFPDSFIDSGEIDVGRASQKIPPGTFWBSOVFIDHPVHLK 465
Db 369 -----DNGKLGFTTHENNTIDSGELDIGRRALQEVPPGIFWRSQFLIDOPQLK 417
QY 466 FNVSLGKAALVGIYGRKGLPPSHTQDFVELLDGRLLTQEARSLLEGTPROSGRTVPSPSS 525
Db 418 FNIQLKDALIGVYGRKGLPPSHTQYDFVELLDGRLLTQEARSLLEGTPROSGRTVPSPSS 477
QY 526 HETGFIQYLDGSIWHLAFYNDGKESEVVSFLTTAIESVDNCPNSCYNGDCISGTCHCFL 585
Db 478 HEAGFIQYLDGSIWHLAFYNDGKESEVVSFLTTAIESVDNCPNSCYNGDCISGTCHCFL 537
QY 586 GFLGPDGCRASCPVLCSGNGQYMKGRCLCHSGKGAECDDVPTNQCIDIVACSNHGTCTGT 645
Db 538 GFLGPDGCRASCPVLCSGNGQYMKGRCLCHSGKGAECDDVPTNQCIDIVACSNHGTCTGT 597
QY 646 CFCNPGYKGESCEEVDCMDPTCSGRGVCVRGSCHECFVWGNGTNCETPRATCLDQCSGHGT 705
Db 598 CACNSGYKGENCEEADCLDPGCSNHGVTGHGCHCNPGWGGSGNCEILKTMCAQDCSGHGT 657
QY 706 FLPTDGLCSDPSWTGHDCSTIEICADCGGHGVGVGTGTCRCEDGMMGAACDORACHPRCA 765
Db 658 YIQESGSCCTCDPNWTGPDCCSNEICVDCSGHGVCMGGSCRCCEGTGTGACNORACHPRCA 717
QY 766 EHGTCRDGKCEGSPGWNGEHCCTIAHLYDRVVKEGCPGLCNGNGRCTLDLNGHCVCQLGW 825

```

[illegible]

Qy	1905	MEYDQAGRIITSRIPADQKTSWYTYLEKSMVLLLHSQOQYIFEPDKNDRLSSVTHMPNVARQ	1964
Db	1857	YDQSQIRIIVSRVFADQKTSWYTYLEKSMVLLLHSQOQYIFEYDMWDRLSAITMPSVARH	1916
Qy	1965	TLETIRSGVYRNIIYQPPENASVIOPTEDGHLHFTYLGTRGRVLYKYKGLSKLAETL	2024
Db	1917	TMQIRISGIYRNIIYNPESNASIITDYNHEGLLQTAFLGTSRRVLKYRQTRLSEIL	1976
Qy	2025	YDTTKVSYFTDETAGMLKTLINLQNEGFTCTIRYQIGPLIDRQIFRTEBGMVNAREFDYN	2084
Db	1977	YDSTRVSFTYDETAGVLKTLNLSQDGFICTIRYQIGPLIDRQIFRESEDCGMVNAREFDS	2036
Qy	2085	YDINSFRVTSQCAVINETPLPIDLYRYDDVSGKTEQFGKFGVYYDINQIILITAWMTTKH	2144
Db	2037	YDINSFRVTSQGVINETPLPIDLYQFPDIISGKVEQFGKFGVYYDINQIILITAWMTTKH	2096
Qy	2145	FDAYGRMKYQEYIEFRSLMTWMTVOYDNMGRVVKKELKVGFPYANTIRSYSEYDADGLOQT	2204
Db	2097	FDAGRIKEIQEYIEFRSLMWITIIQYDNMGRVTKREIKIGFPANTIRKYAYEYDVGLOQT	2156
Qy	2205	YSINDKPLWYSYDNLNGNLHLSPGNSARLTPURYDIRITRLDGVQYXWDEBGFQROR	2264
Db	2157	YYLNEKIMWRYNYDNLNGNLHLNLPSSARLTPURYDIRITRLDGVQYXWDEBGFQROR	2216
Qy	2265	GGDIFEYNSAGLLIKAYNRAGSVSVRYDGLGRVSSKSSHHLHFOFFVADULTNPTKVT	2324
Db	2217	GTEIFEYSSKGLLTRYVYSKSGSWTVIYDGLGRVSSKTSLQOHLFOFFVADULTPYTRIT	2276
Qy	2325	HLYNHSSSEITSYDYLQGLHFAWELSSGBEFYIACNITPLAVFSGTGLMTIKQLIYTA	2384
Db	2277	HVYNHSSSEITSYDYLQGLHFAWEISSGBEFYIASDNTGTPLAVFSSNGMLKQTOYTA	2336
Qy	2385	YGEIYDNTNPNFOLIIIGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHXLKHLSSNV	2444
Db	2337	YGEIYFDSNVDFQVLVGFHGGLYDPLTKLVHGERDYDLAGRWTTPDIETWKRI-GKOP	2395
Qy	2445	MPNLNFKNNNPISNSQDIKCEMTDYNMULLTFGLQHNVI PGYPKPDMDAMEPSYELI	2504
Db	2396	APNLNMYFRNNNPASKIHDVKDITVDNMSLVTFGFLHNVAI PGFPVPKFDLTPSYELV	2455
Qy	2505	HTQWKTQEWDNSKSLIGVQCEVQKQKAPVTLBERFDOLYGSTITSQOQAPKTK- - - - -	2558
Db	2456	-----KSOQWEDVPPIIFGVQOQVARQAKAFUSLQMAEV- - - - -QVSERRKAGAEQSW	2502
Qy	2559	--FASSGSVEGKVKFALKDGRVTIIDISVANEDGRRVAAIILNHAHYLENLHFTIDGVDT	2616
Db	2503	LWFATVKSLLIGKVMILAVSQRVOTVNLINANEDCIKVAALNNAFYLENLHFTIEGKDT	2562
Qy	2617	HYFVKPSPSEGDIAIIGLSGGRTRLENGVNTVVSQINTVLNGTRTRYDTIDLOQYGALCLN	2676
Db	2563	HYFIKTTPESDIGTLURLTSGRKALENGINTVVSQTTVYNGTRTRFADYEMQFGALAH	2622
Qy	2677	TRYGTTLDDEKARVLELARAQVQAWAREQQBLRSGEGLRAWTEGEKQOVLSTGRVQG	2736
Db	2623	VRIGMTLDDEKARILQARQARALARAWAREQQAVRQGEGARLWTEGEKQILLSAKVQG	2682
Qy	2737	YDGFVVISVEQYPELSDSANNIHFMSQSEMGRR	2769
Db	2683	YDGYVYLSVEQYPELADSANNI OFLRSQSEIGKR	2715

```

RESULT 8
US-10-029-020-51
; Sequence 51, Application US/10029020
; Publication No. US20040033971A1
; GENERAL INFORMATION:
; APPLICANT: Gangolli et al.
; TITLE OF INVENTION: Polypeptides and Nucleic Acids Encoding Same
; FILE REFERENCE: 21403-225
; CURRENT APPLICATION NUMBER: US/10/029,020
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,704

```

```

; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/311,590
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/257,314
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 60/311,613
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/315,617
; PRIOR FILING DATE: 2001-08-29
; PRIOR APPLICATION NUMBER: 60/307,506
; PRIOR FILING DATE: 2001-07-24
; PRIOR APPLICATION NUMBER: 60/322,358
; PRIOR FILING DATE: 2001-09-14
; PRIOR APPLICATION NUMBER: 60/294,075
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/288,153
; PRIOR FILING DATE: 2001-05-02
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 51
; LENGTH: 2715
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-029-020-51

```

```

Query Match      68.7%; Score 10227; DB 12; Length 2715;
Best Local Similarity 66.7%; Pred. No. 0;
Matches 1864; Conservative 370; Mismatches 457; Indels 102; Gaps 19;

```

```

Qy 1 MDVKRPYSLTR-RADAERRTSSADSEEGKAP-QKSYSSSEILKAYDQD-ARLAYG 57
Db 1 MDVKRRPYSLTKSRREKERRYNSADNEECRVPTQKSYSSSEILKAFDHDYRLLYG 60
Qy 58 SRVKDIPVQAEFCRTGANTLRELGLBEVTPPHGTLYRTDGLPCQGYSMGAGSDADM 117
Db 61 NRKDLVHRADYTRQGNFTLRQLGVCESATRGVAFCAENGLPHRGYSISAGSDAT 120
Qy 118 EADTVLSPEHPVRLWGRSTRSGSSCLSSRANSLTLTDEHEN---TETDHPGLQNEA 174
Db 121 ENAVMPEFAMRLWGRVSGRSGSSCLSSRSNSALTLTDEHENRSDSESEQSNPPGP 180
Qy 175 RLRTPPPLSHAHTPNQOHAASINSLNRGNFTPRSPSPATPHSLSGEPAGGAQ--- 230
Db 181 TLQPLPP--SHKQHPAQHH--PSTLSLNRSLTNRNQSPAP-----PAALPASELQT 229
Qy 231 -EPAHAQENLLNSNIPLETRNLKQPLGLTQDNLEMDILGASRHGDAYSDGHLFXP 289
Db 230 PESVLQDSNVLSGNVPLESR-----HELFKT 256
Qy 290 G-GTSPFLCTTSPOYPLTSTTVISPPRPLPRSTFAPPAENLKKPSKYCWKCAALSAIV 348
Db 257 GTGTPPLFSTATPGYTWASGVSPPTPLPRLNLTLSRAFAFKSKSKYCSWRCTALCAVG 316
Qy 349 ISATLVILLAVFAMHLFGLNHLQHPQWQWYE---ITEDTASSWPVPTDVSLYPSGGTG 405
Db 317 VSVLLAILLSYFIAMHLFGLNHLQHTENTDFENGKNSDT-----VPTNTVSLPSG--- 368
Qy 406 LETPDRKGKGTGKPSFPFEDSFDISGEIDVGRASQKIPPGTFWRSQVFDHPVHLK 465
Db 369 -----DNKLGFTHENNITDSGELDIGRAIQEVPPGIFWRSOLFIDQPFKLK 417
Qy 466 FNVSLGYAALVGIYGRKGLPPSHQTFDFVLLDGRLLTOEARSLGTEPQSRTGTVPPSS 525
Db 418 FNIQLQDALIGVYGRKGLPPSHQTFDFVLLDGRLLTOEARSLGTEPQSRTGTVPPSS 477
Qy 526 HETFIQYLDGSIWHLAFYNDGKESVSVFLLTAIESVDNCPNCGYNGDCISGTCCHFL 585
Db 478 HEAGFIQYLDGSIWHLAFYNDGKPEQVSNFTIVIESVVECPNCHNGECVSGTCHCFF 537
Qy 586 GFLGPDGRASCPLVCSNGQYMKGRCLCHSGWKGAECDPVTCQIDVACNSHGTCTGT 645
Db 538 GFLGPDGCSRAACPVLCSNGQYSGRCLCFSGMWKGTGTECDVPTTQCIDPQCGGRGICMGS 597

```

```

Qy 646 CINCNGYKGBESCEBVDMDPTCSGRGVCVRGCHCFVWGGTNCETPRATCLDQCSGHGT 705
Db 598 CACNSGYKGENCEEADCLDPCSSNHGVCIHGECHCNFGWGSNCIELKTMCADQCSGHGT 657
Qy 706 FLPTGLCSDPSSWTHGDCSIEICADCGGHGVCVGTGTCRCEDGMMGAACDQACHPRCA 765
Db 658 YLQESGSGCTCDPNWTGPDSCNEICSVDCGSHGVCMGSGSCRCBEGWTGPACNQACHPRCA 717
Qy 766 BHGTCTGCKGCRCSPGWNGEHCTIAHYLDLVVKEGCPGLCNGNGRCTLDLNGHWCVCOLGW 825
Db 718 BHGTCKGCKGCRCSQWNGEHCTIAHYLDLVVKEGCPGLCNGNGRCTLDLNGHWCVCOPGW 777
Qy 826 RGAGCDTSMETACGDSKNDGDGLVDCMDPCCLOP.LCHINPLCLGSPNPLDIIQETQVP 885
Db 778 RGAGCDVAMETLCTDSKNDGDGLIDCWDPCCLQSSCONQPYCKGLPDPDIIISQSLQT 837
Qy 886 VSOQNLHSFYDRIKFLVGRDSTHII.PGENPFDGGHACVIRGQVMTSGTLVGVNISPVN 945
Db 838 PSQQAASFYDRI8FLIGSDSTHVLPGESPENKSLASVIRGQVLTADGTPGLGVNVSFLH 897
Qy 946 NPLFGYTISRQDGSFDLVNNGIGISIIIRFERAPITQEHHTLWLDPWFFVNMETIIMRHEE 1005
Db 898 YSEGYTITRQDGMFDLVANGASLTLVFERSPLTQVHTVMI.PMNVFFVMDTLVMKKEE 957
Qy 1006 NEIPSCDLSNFARPNPVVSPSLTSFASCAEKGPIVPEIOALQBEISISCKMRSLVLS 1065
Db 958 NDIPSCDLSGFRSPSPIIIVSSPLSTFFRSPSPEDSPIIPEQTQVLTBEETIPDTDLKLSYLS 1017
Qy 1066 SRTPGYKSVLRISLTHPTIPFNLMKVLMAVAVGRLFRKFAAAPDLISYFIWKTQVYN 1125
Db 1018 SRAAGYKSVLKITWQAVIPFNLMKVLMAVAVGRLFRKFAAAPDLISYFIWKTQVYN 1077
Qy 1126 QKVFGLSEAFVSUYEYESCPDLILWEKRTTVLQGYEIDASKLGSWLDKHALNIQSGI 1185
Db 1078 QKVFGLSEAFVSUYEYESCPDLILWEKRTTVLQGYEIDASKLGSWLDKHALNIQSGI 1137
Qy 1186 LHKNGENQFVSQPPVIGSIMGCRRESISCPSCNGLADGCKLLAPVALTCGSDGSLVY 1245
Db 1138 LYKNGENQFVSQPPVIGSIMGCRRESISCPSCNGLADGCKLLAPVALTCGSDGSLVY 1197
Qy 1246 GDFNVRIRIFPSGNVTNILELRNDFRSHSPAHKYLATDPMGSAVFLSDSNRRVPKI 1305
Db 1198 GDFNVRIRIFPSGNVTNILELRNDFRSHSPAHKYLATDPMGSAVFLSDSNRRVPKI 1257
Qy 1306 KSTVVVVDLVKXNSEVAGTGOCLPDDTRCGDGGKATEALITNPRGTVDFGLIYFVD 1365
Db 1258 KSLTGAKDLTKNAEYVAGTGOCLPDDTRCGDGGKATEALITNPRGTVDFGLIYFVD 1317
Qy 1366 GTMIRIRIQNGIITLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLD 1425
Db 1318 GTMIRIRIQNGIITLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLD 1377
Qy 1426 NNVLQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHTALESATALAVSHNGVLYTAE 1485
Db 1378 NNVLQITENRQVRIAAGRPMHCQVPGVE-YVPGVGHAVQTTLESATALAVSYGVLYTAE 1436
Qy 1486 TDEKKNIRQVTTSGEISLVAGAPSGCDKNDANCDFSGDDGKAKAKLTSPSSLAVC 1545
Db 1437 TDEKKNIRQVTTSGEISLVAGAPSGCDKNDANCDFSGDDGKAKAKLTSPSSLAVC 1496
Qy 1546 ADGELYVADLGNIRIRIRKKNKPFNTQNMVELSSPIDQELYLFTTQKHLTQSLPTGD 1605
Db 1497 PDGLTYIADLGNIRIRAVSKNKLPLNSMNFYEASPTQELYIFDINGTHQVTVSLVTGD 1556
Qy 1606 YLYNTYTGDDGDTILITDNGNMVNRDSTGMPLWLVVPGDQVVTWMTGNSALKSVTT 1665
Db 1557 YLYNFSYNDNDVTAVTDSNGNLTLRIRDPNRPVVRSPDNQVITWLTGTCGLKSMTPA 1616
Qy 1666 QGHELMAMTYHNSGLLATKSNENGTWTFEYDTSFGRLTNTVFTPTQGVSSFRSDTSSVH 1725
Db 1617 QGHELMAMTYHNSGLLATKSNENGTWTFEYDTSFGRLTNTVFTPTQGVSSFRSDTSSVH 1676
Qy 1726 VQVETSSK--DDVTITTNLSAGAFVTLQDQVRNSYIIGADGSLRLLLANGMEVALQTEP 1784

```

Db 1677 VDISSREEDVSTISNLSSIDSFTVWQOLNSYIGYDGSIRIFYASGLDSHYQTBP 1736
Qy 1785 HLLAGTNPVTKRNVTLPTDNGNLNVWQRKEQAGQVTVFGRRLRVNRNLLSLDDP 1844
Db 1737 HVLAGTANPTVAKENMTLPBGNGQNLVEMFRKEQAGQVNVFGRKLVRNRLSLVDEP 1796
Qy 1845 RVTRTEKIDHRRKFTLRIILYDOAGRPLWSRSSRLVGNVNTYSPGYIAGIORGWSR 1904
Db 1797 RTTKTEKIDHRRKFLRIAYDSCHFTLWLPSSKLMVNVTVSSGQIASIORGTSSEK 1856
Qy 1905 MEYDOAGRIISRIPADGKTSYTYLEKSMVLLHSQRQYIFEPDKNDRLSSVTMPNARQ 1964
Db 1857 VYDSQGRIVSRVPADGKTSYTYLEKSMVLLHSQRQYIFEYDMWDLSSAITMPSVARH 1916
Qy 1965 TLETIRSVGYRNIYQPEGNASVIOPTEDGHLHTFYLTGTRRVYKYKGLSKLAETL 2024
Db 1917 TMQTIIRSIGYRNIYNPESNASIITDYNBEGLLQTAFLGTSRRVLFKYRQTRLSEIL 1976
Qy 2025 YDTTKVSTYDETAGMLKTINLQNEGFTCTIRYRQIGPLIDRQIFRTEEGMVNARFDYN 2084
Db 1977 YDSRVSFTYDETAGVLTNQLSDGFCITIRYRQIGPLIDRQIFRSEDGMVNAFDFS 2036
Qy 2085 YDNSFRVTSQAVINETPLIDIRYDDVSKTEQFGKGVYIYDINOIITAVMHTKH 2144
Db 2037 YDNSFRVTSQAVINETPLIDIRYDDVSKTEQFGKGVYIYDINOIITAVMHTKH 2096
Qy 2145 FDYGRMKEYQYEIFRSLYMTVQYDNMGRVVKXELKVPYANTRYSEYDADQLOQT 2204
Db 2097 FDAHGRLEKQYELFRSLYMTVQYDNMGRVVKXELKVPYANTRYSEYDADQLOQT 2156
Qy 2205 VSINDKPLWYSYDLNGLHLLSPGNSARLTPLRYDIRDITRLGQVQYXMDDEGFLQR 2264
Db 2157 VYLNKIMWRYNDLNGNLHLLSPGNSARLTPLRYDIRDITRLGQVQYXMDDEGFLQR 2216
Qy 2265 GGDIFENYNSAGLILKAYNRAGNSVRYRVDGLGRRYSSKSHHLLQFFYADLTNPTKYT 2324
Db 2217 GTLFEFSSKGLTRVYKSGSGMTVYRYDGLGRRYSSKSHHLLQFFYADLTNPTKYT 2276
Qy 2325 HLYNHSSEITSLYDLQHLFAMELSSGDEFFIACDNTGTLPLAVFSGTGLMIKQILYTA 2384
Db 2277 HLYNHSSEITSLYDLQHLFAMELSSGDEFFIACDNTGTLPLAVFSSNGLMLKQIYTA 2336
Qy 2385 YGELYMDTNFQIIGYHGGYDPLTKLVHMRDYYDLVLAGRWTSFDPHMLKXHLSSNV 2444
Db 2337 YGELYMDTNFQIIGYHGGYDPLTKLVHMRDYYDLVLAGRWTSFDPHMLKXHLSSNV 2395
Qy 2445 MPFNLYMFKNNPISNSQDIKCFMTDVNSWLLTFFGQLHNVPYGPYKPDMDAMERSYELI 2504
Db 2396 APFNLYMFRNNPASKHVDKDIITDVNSWLVTFEGHNLNAFGPPVPKFDLTSPSYELV 2455
Qy 2505 HTOMKQEWNSKSLIGVQCEVQKQKAFVTLERFPOLYGSTITSQQAPKTKK----- 2558
Db 2456 -----KSOQWEDVPPPIFGVQVQVQARAKAFSLGKMAEV-----QVSRKGAEGOSW 2502
Qy 2559 --PASSGVFGKVKALKDGRVTTDIISVANEDGRVAALINHAHYLENLHFTIDGVD 2616
Db 2503 LWFAVKSLLGKGMVAVSGRVQTVNLAINEDCIKVAALVNLNAYLENLHFTIEGKDT 2562
Qy 2617 HYTVKPGSPGDLAAILGLSGRRLTNGVNVTVSQINTVNGRTRRYTDIQLQYCALCIN 2676
Db 2563 HYPIKTTTPSDIGTLRLTSGKALENGINVTYSQSTTVNGRTRRFADVENQFGALALH 2622
Qy 2677 TRYGTTLDEKARVLELQARVQARVQARVQARVQARVQARVQARVQARVQARVQARV 2736
Db 2623 VRYGMLDEKARVLELQARVQARVQARVQARVQARVQARVQARVQARVQARVQARV 2682
Qy 2737 YDGFVVISVEQYBELSDSANNHFMQSEMGR 2769
Db 2683 YDGYVVISVEQYBELADSANNIQLRQSEIGR 2715

US-10-038-854-38
; Sequence 38, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernet, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shmukets, Richard A
; APPLICANT: Tchernov, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangoli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 38
; LENGTH: 2721
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-038-854-38
Query Match 68.7%; Score 10221; DB 16; Length 2721;
Best Local Similarity 67.0%; Pred. No. 0;
Matches 1873; Conservative 362; Mismatches 461; Indels 98; Gaps 20;

Qy 1 MDVKERPYSLTR-RRDAERRYTSASDEEGKAP-QKSYSSSETLKAYDQD-ASLAYG 57

Db 1 MDVKERPYSLTR-RRDAERRYTSASDEEGKAP-QKSYSSSETLKAYDQD-ASLAYG 60

Qy 58 SRVKDIVPQAEBCFRTGTANFTLRELGLBEVTPPHGLTYRTDGLPQCGYSMGAGSDAM 117


```

QY 2259 GFLRQGRGDIPEYNSAGLLIKAYNRAGSWVRYDGLGRVSSKSSHHLQFFVADLT 2318
Db 2217 GFLRQGRTEIFEYSSKGLLTVYSGSGWTVIYRYDGLGRVSSKSLGQHLQFFVADLT 2276
QY 2319 NPTKTHLYNHSSEIISLYDLOCHLFAMELSGDEFFIACDNIGTPPLAVFGTGMUK 2378
Db 2277 YPTRITHVYHSSSEIISLYDLOCHLFAMEISSGDEFFIASDNVTGTPLAVFSSNGLMK 2336
QY 2379 QILTAYGEIYMDNPNFQIIIGVHGGYDPLTKLVHMGRRDYDVLGRWTSDDHELWKH 2438
Db 2337 QIQYAYGEIYFDSNIDFQVIGFHGGYDPLTKLIHFGRRDYDILAGRWTTFDIELWR 2396
QY 2439 LSSNVMPFNLYMFRNNPNISNDIKCFMTDVSNSWLLTFGQHLNVPYKPKDMDAME 2498
Db 2397 I-GKDPAPFNLYMFRNNPNASKIHVDKDYITDVSNSWLVTFGHLNNAIPGFPVPEKDLTE 2455
QY 2499 PSYELIHTQMTQBWDSKSLTGVQCEVQKQAFVTLERFDQLYGSTTSCQAPKTKK 2558
Db 2456 PSYELV---KSQWDDIPPIFGVQQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 2507
QY 2559 ---PASSGVFGKGVKPFALDKGRVTDDIISVANEDGRRVAAILNHAHYLENHFHTIDGVD 2615
Db 2508 WLWPAFVTKSLIGKWLAVSGRVQVTVNLVNIANEDCIKVAALNNAFYLENHFHTIEGKD 2567
QY 2616 THYFVKPSPGEGDLAILGLSGRRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYCALCL 2675
Db 2568 THYFIKTTTDESIGTURLTSGRKALENGINVTVSQSTTVVNGTRTRFADVENQFQALAL 2627
QY 2676 NTRYGITLDBEKARVLRLARQAVRQAWAREQORLEGEGLRAWTEGKQOVLSTGRVQ 2735
Db 2628 HVRVGMVLDBEKARILBQARALARAWAREQVRDGEGLRWTEGKQQLLSAGKVQ 2687
QY 2736 GYDGFVSVQVPELSDSANNHFWQSEMGRR 2769
Db 2688 GYDGYVLSVEQVPELADSANNIQLFQSEIGRR 2721

```

```

RESULT 10
US-10-038-854-36
; Sequence 36, Application US/10038854
; Publication No. US20040022781A1
; GENERAL INFORMATION:
; APPLICANT: Spytex, Kimberly A
; APPLICANT: Li, Li
; APPLICANT: Wolenc, Adam R
; APPLICANT: Vernat, Corine
; APPLICANT: Eisen, Andrew J
; APPLICANT: Liu, Xiaohong
; APPLICANT: Malyankar, Uriel M
; APPLICANT: Shimkets, Richard A
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spaderna, Steven K
; APPLICANT: Gorman, Linda
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Patturajan, Meera
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Gangolli, Esha A
; APPLICANT: Guo, Xiaojia S
; APPLICANT: Shenoy, Suresh G
; APPLICANT: Rastelli, Luca
; APPLICANT: Casman, Stacie J
; APPLICANT: Boldog, Ferenc
; APPLICANT: Burgess, Catherine E
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Ellerman, Karen
; APPLICANT: Gunther, Erik
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038, 854

```

```

; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2003-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 36
; LENGTH: 2725
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-038-854-36

Query Match 68.5%; Score 10201; DB 16; Length 2725;
Best Local Similarity 66.9%; Pred. No. 0;
Matches 1872; Conservative 362; Mismatches 462; Indels 102; Gaps 21;

QY 1 MDVKERPYRSLTR-RRDAERRYTSADSGEGRAP-QKSYSSSETLKAYDQD-ARLAYG 57
Db 1 MDVKERPYCSLTSKSRREKERRYNSADNEECRVPTQKSYSSSETLKAFDHSSRLLYG 60
QY 58 SRVKDIPQBAEPCRTGANFTRELGLEVTPPHGTYRTDGLPCGYSMGAGSDADM 117
Db 61 NRKDLVHREADEFTROQNFTRLQLGVCEPATRRGLAFCAEMGLPHRGYSISAGSDADT 120
QY 118 EADTVLSPEHPVRLWGRSTRSGSSCLSSRANSNLTLTDTEHEN---TETDHEGGLQNH 174
Db 121 ENEAVMPEHAMRLWGRGVKSGSSCLSSRANSNLTLTDTEHENKSDSENEQFASNOGQS 180
QY 175 RLRTPPPLSHAHTPNQHHASINSLNRGNFTPRSNPSPAPTDSHLSGEPPAGAQ---- 230
Db 181 TLQPLPP--SHKQHSACHH-PSITSLNRNSLUNRRNQSPAP-----PALPAELQTT 229
QY 231 -EPAAQENWLLNSNPLETRNLGKQFPFLGTLODNLIEDILGASHRDGYSCHFLFKP 289
Db 230 PESVOLQDSWLVGNSNPLES-----HFLFKT 256
QY 290 G-GTSPFLCTTSPGYPLTSTVYSPPPRPLPRSTFARPAFNKXPKYCNWKAALSAIV 348
Db 257 GTGTTPLFSTATPGYTWASGVVSPTRPLPRNTLSRSFAKFKKSKYSKWKATCAVVG 316
QY 349 ISATLVILLAYFVAMHLFGLNWHLQPMQWYB---ITDITASWVPVPTDVSLYPSGGTG 405
Db 317 VSVLLAILLSYFIAMHLFGLNWLQLOOTENDTFENGKVNSDT-----MPTNVTSLPSG--- 368
QY 406 LETPDRKGTTEGKPSFPDESFTDSGEIDVGRASOKIPPGTFRWQVTDHHPVHLK 465
Db 369 -----DNGLGGTQENNTDSELDIGRRALQEIIPGTFWRQLFDQOFLK 417
QY 466 FNVSLGKAALVGYGRKGLPPSHT---QDFVELLDGRRLLTQEARSELEGTQSRGTG 521
Db 418 FNISLQKDALIGYGRKGLPPSHTQSSPOVDFVELLDGSRLLIAREQSLLETERAGQAR 477
QY 522 PPSSHETGTFQYLDGSIWHLAFYNDGKESEVVSFLTATLESVDNCPSCNCGNGDCISGTC 581
Db 478 SVSLHEAGFIQYLDGSIWHLAFYNDGKNAEQVSFNIVIESVVECPRNCHNGECVSGTC 537

```

582 HCFGLFGLPCGRASCPLVCSGQYMKGRCLCHSGWKGAECDPVNTQCIDVACSHGTC 641
538 HCPGFLGPCSRAACFVLCSGQYSGKRCCLFCFSGWKGTCDVPTTQCIDPQCGRGIC 597
642 ITGTCINPGVKGESCEVDMDPTCSGRGVCVRGCHCFVGGNGNCEPFRATCLDQCS 701
598 IMGSACSSGYKGSCEEAICDIPGCSNHGVCHGECHCFPGWGGNCEILKTNCFDQCS 657
702 GHGTFELPDTGLCSDDPSWTGDCSIEICAAADCGHGVCGGTCTCRCDGMGAACDORACH 761
658 GHGTYLQESGCTCDPNWTGDCSNEICSVDCSHGVCMGGTCTCRCEBGTGPACNQACH 717
762 PRCAEHGTCRDGCECSPGNHGECHTIAHVLDRV-----KSGCPGLCNGNCRCTLDLN 815
718 PRCAEHGTCRDGCECSPGNHGECHTIAHVLDRV-----KSGCPGLCNGNCRCTLDLN 777
816 GWHCVCOLGWRGACDTSMTACDSDKNDGDGLVDCMDPDCCLQPLCHNPLCLGSPNP 875
778 GGHCVCPGWRGACDVAMETLCTDSKDNESDGLVDCMDPDCCLQSSCQNPYCRGLPDP 837
876 LDIIQETQVPVQONLHSFYDRIKFLVGRDSTHIIPOENPFDPGGHACVIRGQVMTSDGTP 935
838 QDIISQSLQSPQQAASFYDRIISFLIGSDSTHVIPGESPFNKSASVIRGQVLTADGTP 897
936 LVGTNISFVANPLFGYTIISQDGSFDLVTNGGISIIILREERAPLITOEHTLMLPWDRFFV 995
898 LIGVNVSEFFHPETGYTIITQDGFDLVANGASLTIVFERSPLTQYHVVWPNVYV 957
996 METIIMRHEENIEPSCDLSFARNPVVSPSLTSPASSCAEKGPVPEIQALQBEISIS 1055
958 MDTLMEKEENDIPSCDLSGVRPNPIIVSPLSTFFRSPPEDSPIPETQVLHEETIP 1017
1056 GCKVRLSYLSRTPGYKSVLRIISLTHPTIPNLKMLHMAVEGRIPKMFAPAAPDLSY 1115
1018 GTDLKLSLSSRAAGYKSVLKITMTQSIIPNLKMLHMAVAVGRLEFQKFPFASPNIAYT 1077
1116 FIWKTIVYQKVGLEAFVSGVYESCPDLILWEKRTVTLQGYEIDASKLGMSLDK 1175
1078 FIWKTDAYNQVYGLSEAVVSGVYESCLDLTLEKRTALLOQYELDASNMGWTLDK 1137
1176 HHALNIOGILHKGNGENQFVSQOPPVIGSMGNRRRSISCPSCNGLADGNKLLAPVAL 1235
1138 HVLVDVQNGILYKNGENQFVSQOPPVVSSIMGNRRRSISCPSCNGADGNKLLAPVAL 1197
1236 TCGSDGSIYVDENYIRRIIPSGNVNTILELRNDFRHSHPAHKYLATDPMGAVFLS 1295
1198 ACGIDGSLYVDGFNVRRIIPSGNVTSLVLELRNDFRHSNPAHYLATDPTVGDLYVS 1257
1296 DSNRRVFKISTVYVVDLVKNSEVAGTGCQCLPDDTRCGDGGKATEATLNPGRITV 1355
1258 DNTNRRIVRPSLTKAKDLTKNAEVVAGTGCQCLPDEARCGDGGKATEATLMSPKGMV 1317
1356 DXFGLIYVDGTMIRRDQNGIISTLLGNDLTGSARPLSCDSVMDISQVLEWPTDLAIN 1415
1318 DKNGLIYVDGTMIRKVDQNGIISTLLGNDLTGSARPLTCDTSMHISQVLEWPTDLAIN 1377
1416 PWDNSLYVDNNVLIQISENHQVRIVAGRPMHCQVPGIDHPLLSKVAIHATLESATALV 1475
1378 PWDNSIYVDNNVLIQITENQVRIIAGRPMHCQVPGVE-YPVGKHAVQITLESATAIV 1436
1476 SHNGVLYTAETDEKKNIRIQVTTSGEISLVAGAPSCDCKNDANCDGFCSDGDAYKADK 1535
1437 SVSGVLYTETDEKKNIRIQVTTDGEISLVAGIPSECDCKNDANCDYOSGDGDAYKADK 1496
1536 LNTPSSLAVCADGELYVADLGNIRIRFRKNKPELNTONMYELSPDQELLYLFTDTGKH 1595
1497 LSAPSSLAASPDGTYIADLGNIRIRAVKKNPLNSMNFVEVASPTDQELIYIIDDINGTH 1556
1596 LYTSQSLPTGDLYNFTYDGDITLITDNGNMVNRDSTGMPLWLVPDQGVYVMTG 1655
1557 QYTVSLVTDGTYLXNFSYNDNDITAVTDSNGNTLRIIRDPNRMPVRVVSVDNQVILTLIG 1616

1656 TNSALKSVTTQGHSLAMVTHGNSGLLATKSNENGWTTTFEYDVSFRLTNTVTEPTGOVSS 1715
1617 TNGCLKMGTAQGLEVLVTFYHNSGLLATKSDETGWITTFDYSSEGRLTNTVTFPGVVTN 1676
1716 FRSDTSDSVHVQVETSSK-DDVTIITNLSAGAFYTLLOQDVNRNSYYICADCSURLLAN 1774
1677 LHGDMDKAITVDIESSREEDVSIITNLSIDSFTYMWQDQLRNSYQIGDGLRIIVAS 1736
1775 GMEVALQTEPHLLAGTVNPTVGRKNTLPIDNLNLVWRORKEQARQVTVFGRDLVH 1834
1737 GLDSHYQTEPHVLAGTANPTVAKRNTLPGENGQNLVWRFRKEQAQKVNVPFGRKLNV 1796
1835 NENLLSLDFDRVTRTEKIYDDHRKFTLILYDQAGRPSLWSPSSRLNGVNTVYSPGGYIA 1894
1797 GRNLLSVDFDRTTEKIYDDHRKFLRLAYDTSCHPTLWLPSSKLMVAVNTYSSTGOIA 1856
1895 GIQRMIGSERMEYDQAGRIITRSIFADGKTWSYTYLEKSMVLLHSHQROYIIFEDKNDRLS 1954
1857 SIQRTGTTSEKVDYDQGRIVSRVAFDAGKTWSYTYLEKSMVLLHSHQROYIIFEYDMWDLRS 1916
1955 SVTMEVNAQTLETITRSYGYRNIYQPPEGNASVIQDTEGHLHTFVLTGTRRVIKY 2014
1917 AITMFSVAHTNQITRSIYGYRNIYNPPESNASIITDNEBGLLQTAFLGTSRVLFKY 1976
2015 GKLSKLAETLYDTTKVSFTYDAGTKITNLQNEGFTCTIRYQIGPLIDRQIFRTEE 2074
1977 RQTRLSEILYDSTRVSTFYDAGTKITNLQSDGFTCTIRYQIGPLIDRQIFRFSBD 2036
2075 GWNARFNDYDNRFRVTSMQAVINETPLDLYDDVSGKTEQFGKGVLYYDINQII 2134
2037 GWNARFNDYDNRFRVTSMQAVINETPLDLYDDVSGKTEQFGKGVLYYDINQII 2096
2135 TTAVMTHTKHFDAYGRMEQVYEIIFRSIMYMTVOYDNNMGRVVKKELKVPYANTTRYX 2194
2097 STAVMTYTHGFDAHGRKEIQEIQEYIFRSIMYMTVOYDNNMGRVVKKELKVPYANTTRYX 2156
2195 EYDADGQLOTVSINKPLWRYSDYDNLNHLSPNSARLTPLRDYDIDRDIRTIRLGDVQYK 2254
2157 EYDADGQLOTVSINKPLWRYSDYDNLNHLSPNSARLTPLRDYDIDRDIRTIRLGDVQYK 2216
2255 MDEDFLROGGDIPEYNSAGLIIKAYNFRAGSWSVRYDYDGLGRVSSKSSSHHLOFFY 2314
2217 LDEDFLROGGDIPEYNSAGLIIKAYNFRAGSWSVRYDYDGLGRVSSKSSSHHLOFFY 2276
2315 ADLTPKTKVTHLYNHSSEITSYLYDLOGLHFLABELSGDEFYIACDNTGTPLAVPSGTG 2374
2277 ADLTPKTKVTHLYNHSSEITSYLYDLOGLHFLABELSGDEFYIACDNTGTPLAVPSGTG 2336
2375 LMKQIYLYTAYGEIYMDTNPFIILYHGGIYDPLTKLVHMGRRDYDVLGRWTSPPDHE 2434
2337 LMKQIYLYTAYGEIYMDTNPFIILYHGGIYDPLTKLVHMGRRDYDVLGRWTSPPDHE 2396
2435 LMKHSSSNVMPFNLYMFKNNNPIINSQDIKCFMTDVNSWLLTFFGQLHNVIPIGPKDM 2494
2397 INKRI-GKDPAPFNLYMFKNNNPIINSQDIKCFMTDVNSWLLTFFGQLHNVIPIGPKDM 2455
2495 DAMEPSYELIHTQMTQEWNDNSKSLGVCEVQKQLKAFVTLERFDOLYGSITSCQAP 2554
2456 DLTEPSYELV---KSCQWDDIPPIFGVQQQVARQAKFLSLGKMAEVQ----VSRRRAG 2507
2555 KTKK---PASSGSVPKGVKFKALKQGRVTTDIISVANEDGRRVAALINHAHLENLHFTI 2611
2508 GAGSWLWPAFVKSILGKQWMAVUSQGRVQTNVINANEDCIKVAALNNAFLENLHFTI 2567
2612 DGVDTHYFKVPSEGLLAILGLSGRRRTLENGVNTVSOINTVNLGRTRRTYDIOLOQV 2671
2568 EGKQTHYFIKTTPTPESDGLTLTSGRKALENGINVTVSQSTTVVNGRTRRFRADVEMQFG 2627
2672 ALCLNTRYGTTLDEEKARVLELAROQAVARQORLECEBEGELRAMTECEKQOVLST 2731
2628 ALAHVRYGTTLDEEKARVLELAROQAVARQORLECEBEGELRAMTECEKQOVLST 2687
2732 GRVQVGDGFVFSVEQYPSLSDSANNIHFMRQSEMR 2769

Db 926 YLSRAAGKSVLKITWTQSIIPENLAKVHLMAVAVGRLFKQFPASPENLAYTFIWDKTD 987
QY 1123 VYNQKVFGLSBAFVSQVGYEBSQDILILWEKRTITVLOGYEIDASKLGGNSLDKXHALNIQ 1182
Db 988 AYNQKVFGLSBAFVSQVGYEBSQDILILWEKRTITVLOGYEIDASKLGGNSLDKXHALNIQ 1047
QY 1183 SGIHLKNGENQFVSQPPVIGSIMGNGRRRSISPCSCNGLADQNKLLAPVALTTCGSDGS 1242
Db 1048 NGILYKNGENQFVSQPPVIGSIMGNGRRRSISPCSCNGLADQNKLLAPVALTTCGSDGS 1107
QY 1243 LYGDENYIRIRPSPGNVTNILEURNKDFRHSHPAHKYYLATDPMGAVFLSDSRRV 1302
Db 1108 LYGDENYIRIRPSPGNVTNILEURNKDFRHSHPAHKYYLATDPMGAVFLSDSRRV 1167
QY 1303 PKIRSTVVKDLVKNSEVAGTGQCLPFDDTRCGDGKATEALTNPRTGIVTDKGLIY 1362
Db 1168 YRPSKLTGAKDLTKNAEVAGTGQCLPFDDTRCGDGKATEALTNPRTGIVTDKGLIY 1227
QY 1363 FVDGTMIRRIDQNGIISTLGSNDLTSARPLSCDSVMDISQVRLWPTDLAINPMDSNY 1422
Db 1228 FVDGTMIRRIDQNGIISTLGSNDLTSARPLSCDSVMDISQVRLWPTDLAINPMDSNY 1287
QY 1423 VLDNNVLOISENHQVIRVAGRMHCQVPGIDHDFLLSKVAIHATLESATALVSHNGVLY 1482
Db 1288 VLDNNVLOISENHQVIRVAGRMHCQVPGIDHDFLLSKVAIHATLESATALVSHNGVLY 1346
QY 1483 IAETDEKKINRIQVTTSGEISLVAGAPSGDCCKNDANCCFSGDDGYADAKLNTSS 1542
Db 1347 IETDEKKINRIQVTTSGEISLVAGAPSGDCCKNDANCCFSGDDGYADAKLNTSS 1406
QY 1543 AVCADGELYVADLNIRIRFRKXKPFINTQNMVELSPIDQELYLFTTGKHLIYQSILP 1602
Db 1407 AASPDGTLIYADLNIRIRFRKXKPFINTQNMVELSPIDQELYLFTTGKHLIYQSILP 1466
QY 1603 TGDYLYNFTYDGDITLITDNGNMVNVRRDSGMPLWLVVDPGVVYVMTGNTSALKS 1662
Db 1467 TGDYLYNFTYDGDITLITDNGNMVNVRRDSGMPLWLVVDPGVVYVMTGNTSALKS 1526
QY 1663 VTTQCHELAMTYHGNGLATKSNENGWTTTFEYSDYFGLRTNVTFTQGVSSFRSDTS 1722
Db 1527 MTAQGLELVFTYHGNGLATKSNENGWTTTFEYSDYFGLRTNVTFTQGVSSFRSDTS 1586
QY 1723 SVHVOVETSSK-DDVTITTNLSASGAFVTLLOQVRSNYIGADGSLRLLLANGMEVALQ 1781
Db 1587 AITVDISSSREEDVSITNSLSSIDSFTYMWQDQNRNSYQIGYDGLRIIYASGLDSHYQ 1646
QY 1782 TEPHLLAGTVNPTVGRKRVNLTPIDNLNLVWRQKQEQVTVFGRRLRVHNRNLLSL 1841
Db 1647 TEPHLLAGTVNPTVGRKRVNLTPIDNLNLVWRQKQEQVTVFGRRLRVHNRNLLSL 1706
QY 1842 DFDVTRTEKTYDHRKFTLILYDQAGRPSLWSPSSRLNGVNTYSPGGVIAGIQRGM 1901
Db 1707 DFDVTRTEKTYDHRKFTLILYDQAGRPSLWSPSSRLNGVNTYSPGGVIAGIQRGM 1766
QY 1902 SERMEYDQAGRTSRI PADGKTWSYTYLEKSMVLLLSORQYIFEFKNDRLSVTTPNV 1961
Db 1767 SERMEYDQAGRTSRI PADGKTWSYTYLEKSMVLLLSORQYIFEFKNDRLSVTTPNV 1826
QY 1962 ARQTLTIRSYGYRNIYQPPGEGNASVIOQDTEPCHLLHTFYLGTGRVRYKYGKLSKLA 2021
Db 1827 ARQTLTIRSYGYRNIYQPPGEGNASVIOQDTEPCHLLHTFYLGTGRVRYKYGKLSKLA 1886
QY 2022 ETLIDYTKVSTYDETAGMLKNTINLONEGFTCTIRYQIGPLIDRQIFRFTTEEGMVARF 2081
Db 1887 ETLIDYTKVSTYDETAGMLKNTINLONEGFTCTIRYQIGPLIDRQIFRFTTEEGMVARF 1946
QY 2082 DYNVNSFRVTSMAQVINETPLPIDLYRYDDVSGKTEQFGKFGVYIYDINOIITAVMTH 2141
Db 1947 DYSYDNSFRVTSMAQVINETPLPIDLYRYDDVSGKTEQFGKFGVYIYDINOIITAVMTH 2006
QY 2142 TGHFDAYGRMEVQVEIFRSLMNTVQYDNNMGRVVKELKVGYPYANTTRYSYDYADGQ 2201

Db 2007 TXHFDAGRIKEIQVEIFRSLMNTVQYDNNMGRVVKELKVGYPYANTTRYSYDYADGQ 2066
QY 2202 LOTVINSNDKPLWRYSYDNLGNLHLLSPGNSARLTPRLYDIDRDIRITRLGDVQVYEMDEGFL 2261
Db 2067 LOTVINSNDKPLWRYSYDNLGNLHLLSPGNSARLTPRLYDIDRDIRITRLGDVQVYEMDEGFL 2126
QY 2262 RQGGDIIFEYNSAGLLIKAYNRAGSVRYRYDGLGRVSSKSHHLOFFYADLTNPT 2321
Db 2127 RQGGDIIFEYNSAGLLIKAYNRAGSVRYRYDGLGRVSSKSHHLOFFYADLTNPT 2186
QY 2322 KYTHLYNHSSSEITSLYDYLQHLFAMELSSGDEFIACDNLGTPLAVPSGTGLMKQL 2381
Db 2187 RITHVYNHSSSEITSLYDYLQHLFAMELSSGDEFIACDNLGTPLAVPSGTGLMKQL 2246
QY 2382 YTAIGEIMYDTPNPFQIIIGYHGGLYDPLTKLVHMRDRDYDVLAGRWTSPDHMLKHLSS 2441
Db 2247 YTAIGEIMYDTPNPFQIIIGYHGGLYDPLTKLVHMRDRDYDVLAGRWTSPDHMLKHLSS 2305
QY 2442 SNVMPFNLYMFKNNPINSQDIKCFMTDVNSWLTTFQFOLHNVIPGYKPKDMDAMEPSY 2501
Db 2306 KDPAPFNLYMFKNNPINSQDIKCFMTDVNSWLTTFQFOLHNVIPGYKPKDMDAMEPSY 2365
QY 2502 ELIHTCMKTOEDNKGSKILGVQCEYQKOLKAPVTLERFDQLYGSTITSQQAPKTK--- 2558
Db 2366 ELV----KXQWDDIPPIFGVQOQVARQAKAFSLGKMAEVQ----VSRERRAGGAQSMWL 2417
QY 2559 FASSGSVFGKGYKPAKDKGRVTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHY 2618
Db 2418 FATVKSLLIGKGYMLAVSQGRVQTNVLNANEDCIKVAALNNAFYLENLHFTIEGKDTY 2477
QY 2619 FVKPSPSEGLAILGLSGGRTLENGVNVTVSQTINVLNGRTRVTDIOLQXGALCLNTR 2678
Db 2478 FIKTTTPESDGLTSLTSGKALENGINVTVSQSTVNVNGRTRRTRRTRRTRRTRRTRRTRR 2537
QY 2679 YGTTLDEEKARVLELQARQAVARQAWARQOORLEGEGLRAWTEGEKQOVLSTGRVQGYD 2738
Db 2538 YGTTLDEEKARVLELQARQAVARQAWARQOORLEGEGLRAWTEGEKQOVLSTGRVQGYD 2597
QY 2739 GFVLSVVEQYPELSDSANNIHFMRQSEMR 2769
Db 2598 GFVLSVVEQYPELSDSANNIHFMRQSEMR 2628

RESULT 12

US-10-038-854-42
; Sequence 42, Application US/10038854
; Publication No. US20040022781A1

GENERAL INFORMATION:

APPLICANT: Spytex, Kimberly A

APPLICANT: Li, Li

APPLICANT: Wolenc, Adam R

APPLICANT: Vernet, Corine

APPLICANT: Eisen, Andrew J

APPLICANT: Liu, Xiaohong

APPLICANT: Malyankar, Uriel M

APPLICANT: Shimkets, Richard A

APPLICANT: Tchernev, Velizar

APPLICANT: Spaderna, Steven K

APPLICANT: Gorman, Linda

APPLICANT: Kekuda, Ramesh

APPLICANT: Patturajan, Meera

APPLICANT: Gusev, Vladimir Y

APPLICANT: Gangolli, Esha A

APPLICANT: Guo, Xiaojia S

APPLICANT: Shenoy, Suresh G

APPLICANT: Casman, Stacie J

APPLICANT: Boldog, Ferenc

APPLICANT: Burgess, Catherine E

APPLICANT: Edinger, Shlomit R

APPLICANT: Ellerman, Karen

APPLICANT: Gunther, Erik

APPLICANT: Smithson, Glennda

328	Db	SLQKDALIGVYGRKLPSPHSTQYDFVLLDGSRLIAREQSKSLLETBRAGQARSVSLHEA	387
529	Qy	GFIQYLDGSIWHLAFYNDKESVVSFLTIAESVDNCPNCSYNGDNCISGTCGHCFPLAPL	588
388	Db	GFIQYLDGSIWHLAFYNDKNAEQVSNTIVIESVVECPNCHGNGECVSGTCHCFPGPL	447
589	Qy	GPDCGRASCPVLCSNGGOYMKGRCLCHSGMKGAECDPVINCOCIDVACSNHGTCITGTCLC	648
448	Db	GPDCSRACPVLCSSNGGYSAGRCLCFSGWKGTEDVPTTCQIDPQCGGRGICIMGSCAC	507
649	Qy	NPYKGXGSEEDVCMDDPFTCSRGVGVYAGECHCFVGMGTCNETPRATCLDQCSGHGTFLP	708
508	Db	NSGYKGSECEADCDIPGCSNHYGCIHGECHSPGMGGSNCEILTKMTCPDQCSGHGTVLQ	567
709	Qy	DTGLCSCPSPWTGHDCSIETICAAACGGHGVYCVGTCRCEDGMGAACDORACHPCAEBH	768
568	Db	EGSCTCDPNTWTPDCSNEICSDVCSGSHGVCMGTNRCBEGWTPCTCNQACHPCAEBH	627
769	Qy	TRDRKCBSCSPGNGEHCHTIAHYLDRVKVGCGPGLCNNGRCPTLDLNGWHVCQJGWRGA	828
628	Db	TKCDKCBKCSHGNGEHCTI-----EGCPGLCNNGRCPTLDQNGWHVCVQSPGWRGA	678
829	Qy	GCDTSMETACGDSKDNDDGSLVDCMDPCCLQLCHINPLCLGSPNPLDIIQETQVPVSQ	888
679	Db	GCDVAMETLCTDSKDNEDGLIIDCMDFDCCJLQSSCQNPYCRGLPDPQDIISQSLQSPSQ	738
889	Qy	QNLHSPFYDRIKFLVGRDSTHIIPEGNPFDGGHACVIRGQVMTSDGTPLGVNNISFVNPL	948
739	Db	QAQKSFYDRIISFLIGSDSTHVIPEGSEPNKSLASVIRGOVLTAGDTPLGVNVSPFFHPE	798
949	Qy	FGYITISROGSDFIATVNTGGISIIILFERAPFTIQEHTLMLPMDRFPVMEIIMRHEENEI	1008
799	Db	YGYTITROQGMFDLVANGGASLTUVPERSPFLTQYHTVMPNVVYFMDTIVMEKEENDI	858

1009	QY	PCDUSNFA	RNPVVS	SPUTS	SSASS	CAEKP	VP	EQALQ	BEI	SISG	CKML	SYLS	SRT	1068
859	Db	PCSDLSG	RVPNP	IV	SP	LS	PT	FR	SP	ED	SP	IP	TQ	VL
1069	QY	PGYKSV	LAI	ST	HT	IP	FN	LK	MV	AV	BGR	L	FR	KW
919	Db	AGYKSV	LAI	TQ	SI	IP	FN	LK	MV	AV	VGR	U	FR	KW
1129	QY	FGLS	EAF	V	S	G	Y	E	S	C	P	D	L	W
979	Db	YGLS	EA	V	S	G	Y	E	S	C	D	L	T	W
1189	QY	NGEN	OF	S	Q	OP	P	V	I	G	S	M	G	N
1039	Db	NGEN	Q	F	I	S	Q	OP	P	V	S	M	G	N
1249	QY	NYIR	RI	PS	G	N	V	T	NI	L	R	N	K	D
1099	Db	NVYR	RI	PS	G	N	V	T	NI	L	R	N	K	D
1309	QY	VVVK	D	L	V	K	S	E	V	A	G	T	G	D
1159	Db	TGAK	D	L	T	K	N	A	E	V	A	G	T	G
1369	QY	RRRI	D	Q	N	G	I	S	T	L	G	S	N	D
1219	Db	IRK	V	D	Q	N	G	I	S	T	L	G	S	N
1429	QY	VLOI	S	E	N	H	Q	V	R	I	V	A	G	R
1279	Db	VLOI	T	E	N	R	Q	V	R	I	V	A	G	R
1489	QY	KKIN	R	I	R	O	V	T	T	S	G	B	I	S
1338	Db	KKIN	R	I	R	O	V	T	T	S	G	B	I	S
1549	QY	EY	V	A	D	L	G	N	I	R	I	P	I	R
1398	Db	TLY	A	D	L	G	N	I	R	I	P	I	R	I

Db 2555 QOTTERHNOAFMALE-----GQVISKRLHASIREKAGHWFATSTPIIGKIMFAVKGR 2608
Qy 2579 VTTDILISVANEDGRRVAAILNHAYLENLHFTTIDGVDTHYFVPGPSEGDAILGLSGGR 2638
Db 2609 VTTGISIAITDDSRKIASVLNSAHYLEKMHYSIEGKDTHTYFVKGADSDLVTLAMTSGR 2668
Qy 2639 RTLENGVNVTVSOLNTVLNCRTRYTDIOLOQALCLNTRYG---TTLDEEKARVLELAR 2695
Db 2669 KVLDSGVNVTVSOTLLINGTRFRFNIEFOYSTLLINTRYGLTATLDEEKARVLDQAR 2728
Qy 2696 ORAVRQWALEQORLREGESGLRAWTEGEKQOVLSTGRVGYDGFVVISVEQYPELSDSA 2755
Db 2729 QRALGSAAWEKQOKARDGREGSRVWTDGKQQLNTGRVGYGYVVLVPEQYPELADSS 2788
Qy 2756 NNIHEVREQSEMGR 2769
Db 2789 SNIQFLQONEMGR 2802

RESULT 14
US-09-800-198-69
; Sequence 69, Application US/09800198
; Publication No. US20030087816A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Cornie AM
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herimann, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishna
; APPLICANT: Mezes, Peter S
; APPLICANT: Rastelli, Luca
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-697
; CURRENT APPLICATION NUMBER: US/09/800,198
; CURRENT FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 69
; LENGTH: 2802
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-800-198-69

Query Match 64.6%; Score 9616.5; DB 10; Length 2802;
Best Local Similarity 61.3%; Pred. No. 0;
Matches 1775; Conservative 411; Mismatches 491; Indels 217; Gaps 27;
Qy 1 MDVKERKPYSLTRRR-DAERRYTSSADSEEGKAP-QKSYSSSETLKAYDQARLAYGS 59
Db 1 MDIKDRR-HRSLTRGRCCKEYRTSSSLDSEDQVPAQKSYSSSETLKAYGHDTRMHYGN 59
Qy 59 RVKDIVPQEAEEFCSTANTLRELGLVEYTPPHGTLRYTDIGLPQCGYSGWAGSDADME 118
Db 60 RVSDLVAREDEFPFQGTNTLTALGL:CBPS-PHRSYCSDDIGLHCGYSLSTGSDADSD 118
Qy 119 ADTVLSPHVPVLRWGRSTRSGRSCSLSSRANSNLTLDTEHEN----- 161
Db 119 TEGGWSPEHAIRLWGRIGKSSRSGSLSSRENSALTLDSDNENKSDENDFPHLSEKLG 178
Qy 162 -----TETH----- 166
Db 179 DROTSWOQLAETKNSLIRPIPTSSSLLPSAQLPSSHNPVPSVQMPLLDNTSHQIM 238
Qy 167 -----PGQLQNH---ASLRTP-PPPLSHAHTPNQHAASI 197
Db 239 DTNPDEEFSPNSYLLRACSGPQOASSGPPNHHQSQTSLRPLPPPHNLSL-----HHSSA 295
Qy 198 NSLMRGNFTPRSN-----PSAPTDLHSLSGBPAGGAQEPAHAGENMLLNINPLETRNLG 253

Db 296 NSLNRSNLTNRNRNQHAPAPAPND--LATTP-----ESVQLQDSWVLNSVPLETR--- 344
Qy 294 KQPLGCTIQDNLIEWDILGASRHDGAYSDCHFLPK-PGTSPLCTTSPGYPLFSSSTVYS 312
Db 345 -----HFLKPTSSGTTPLFSSSSPGPYLSTGTYT 374
Qy 313 PPPRPLRSTFARPAFNLKPKSKYCNWKAALSIAIVSATLVLILAYVAMHLEGLNWHL 372
Db 375 PPPRLLPNTFSRNFAPLKKPKSKYCNWKAALSIAIAAVALLAILLAYFIAMHLLGLNWL 434
Qy 373 QPMEQMVEI-----TEDTASWVPVPTDVSLYPSGGTGLTPDRKGKGTTEGKPSF 424
Db 435 QPADGHTFESNGLRPGAGAAGDAAAPPA-----GRG-----PW 467
Qy 425 PEDFDIISGIDVGRASQKIPPTFWRSOVFIIDHPVHLKFNVSIGKAAALVGIYGRKL 484
Db 468 VTRNSIDSGTEVGRKVTQEPFGVFWRSQIHISQOFLAFKFNISLGDALFGVYIRGL 527
Qy 485 PPSHTQFDFVELLDGRRLLTQEARLEGTTPQSGRTVPSPSHSETGFIQYLDGSIWHLAFY 544
Db 528 PPSHAQYDFMERLDGK-----EKMSVWESPERRSIQTLVQNEAVFVQYLDVGLMHLAFY 582
Qy 545 NDGKESEVVSLTITAIESVDNCPNCGNGCISGTCFGLGFGDPCGRASCVPVLCGN 604
Db 583 NDGKDEKVVSTVLDSVQDCPRNCHNGEGCVGCHCFPGFGADCAKAAACVPVLCGN 642
Qy 605 GQMKGRCLCHSGMKGAECVPTNQCIDVACSNHGTCTGTICINPGYKGBSCSEVDCMD 664
Db 643 GQYSGTCLCYSGWKGPEDVPISQIDPSCGGHSGCIEGNCVCSIGYKGENCEBVDCLD 702
Qy 655 PTCGRGVGVGECHCFVWGVTNCTPRATCLDQCSHGCTFLPDTGLSCDPSWTGHDG 724
Db 703 PTCNHHGVGVNGECLCSPGWGGINCELPRACQPCQSGHGYLSLDTGLSCDPMWMPDC 762
Qy 725 STEICAAADCGGHGVGVGTCTCEDGWMGAAADQACHPRCAEHGTCRDKGCEBSPWNGE 784
Db 763 SVEVCSVDCGTHGVGIGACRCEGWTGACDQVCHRCRTEHGTCKDGKCECRGNGE 822
Qy 785 HCTIAHYLDVVYKGCPCGLCNGNGRGTLDLNGWHCVQGLGWRGAGCDSMETACDSKDN 844
Db 823 HCTIGRQTGTETDGCPLCNGRGTLCQNSWCQVCQCTGWRGPGCNVAMETSCADNKN 882
Qy 845 DGDGLVDCWDPDCCLOPLCHINPLCLGSPNPLDIIQETQVPSVQOQLHSHFYDRIKFLVGR 904
Db 883 EGDGLVDCVLPDCCLOSTCQNSLLCRGSRDPLDIIQQSH--SGSPAVKSFYDRIKLLVGR 940
Qy 905 DSTHIIPGENPFDGGHACVIRGQVMTSDGTPLGVNISFVNNPLFGYITISRQDGFDLVT 964
Db 941 DSTHIIPGENPNSLSLIRGQVTTDGTPLGVNVSFKYKYGITRQDGMFDLVA 1000
Qy 965 NGGIIILIRFERAPITQEHILWLPWDRFFVWETIIMRHEENIEPSCDLSNFARNPVYS 1024
Db 1001 NGGSSLTLLHFERAPFMSQERTVWLPWNSFYAMDTLVKMTENSIPSCDLSGVRDPPVII 1060
Qy 1025 PPLTSFASSCAEKGPVPEIQALQEEISISGCKMRLSYLSRTPGYKSVLRISLTHPTI 1084
Db 1061 SPSLTFSDAPGRNPIVPTQVLEHEIEVFGSSIKLYLSRRTAGYKSLKLIINTQSLV 1120
Qy 1085 PNLKMKVLMVAVEGRILFRKWFAPADLSYFPIWDTQVYNQKVPGLSEAFVSVGYEYES 1144
Db 1121 PNLKIKVLMVAVEGHLFKGFLASPNLAYTFINDKTDAYGQKYGCLSDAVVSVGFYET 1180
Qy 1145 CPDLILWEKRTVLOGYEIDASKLGWSLDKXHALNIOGILHKGNGENQFVSQQPPVIG 1204
Db 1181 CSLLILWEKRTALQGFELDPNLGWSLDDKHVNLNKSILHKGNGENQFUTQPAVIT 1240
Qy 1205 SIMNGRRRSISPCPCNGLADGNKLLAPVALTCGSDGLYVGDFFNYIRRIIPSGNVNIL 1264
Db 1241 SIMNGRRRSISPCPCNGLAEGNKLLAPVALAVGIDGSLFVGDFNYIRRIIPSRNVTSL 1300
Qy 1265 ELRNKDFRSHSPAKHYLATDPMGAVFLSDNSRVRFKISTVTVVLDLVKNSVWAGT 1324
Db 1301 ELRNKEFKHNSNPAKHYLAVDVPSGSLYVSDTNRRRIYKVKSLTGTDXDLAGNSVWAGT 1360

QY 1325 GDOCLPFDTRCGGKATBATLNPRTGIVTDKFLIYVDGTMIRRIDONGIISTLLGS 1384
Db 1361 GEQCLPDEARCGGGKAVATLMSPRGIAVKYGLMYFDATMIRKVDQNGIISTLLGS 1420
QY 1385 NDLTSARPLSCDSVMDISQVRLEWPTDLAINPDMSLYLDNNVYLOISNHOVRIVAGR 1444
Db 1421 NDLTAVRPLSCDSVMDISQVRLEWPTDLAVPMDNSLYVLENNVILRTITENHOVSIAGR 1480
QY 1445 PMHCQVPGIDHFLKVAIHATLESATALAVSNGVLYIAETDEKKNIRIROVTTSGEIS 1504
Db 1481 PMHCQVPGID-YSLSUKAIHSALESAGIAISHTGVLYISETDEKKNIRIROVTTSGEIS 1539
QY 1505 LVAGAPGCCCKNDANCDGSDGYAKDAKLATPSSLAVCADGELYVADLGNIRIRFIR 1564
Db 1540 LLAGAASDCCKNDVNCYSGDDGYATDAILNPSPLAVAPDGT-VIADLGNIRIRAVS 1599
QY 1565 KXKFFLNTQMYELSSPIDOELYLFDTTGKHLTQSLPTGDIYLYNTYTCDDGITLITDN 1624
Db 1600 KXRPILNSFNQYEAAASPGOEYVFNADGTHQVTLVLTGEXLYNFTYSSDNDVTEWDS 1659
QY 1625 NGNMVNRDSTGMPLMLVVPDGOVYVMTGNTSALKSVTTQGHLEAMTYHGNISGLAT 1684
Db 1660 NGNSLVKRRDASGNPHLLMPDQIVTLAVGTNGGLKLVSTQTLGLMTYNGNSGLAT 1719
QY 1685 KSNENGWTFYEDSFGRLTNVPTPTQOVSSFRSDTSSVHVQVETSSK-DDVITITNLS 1743
Db 1720 KSDGTGWTTFYDHEGLTNVTRPTGVVTSLHREMEKSTIIDIENSRRDDVVITNLS 1779
QY 1744 ASGAFYTLLODOVNSYIYGADGSLRLLLANGHEVALQTEPHLLAGTVNPTVGRNVLP 1803
Db 1780 SVEASYTVQOVVNSVQLCNGTLRVYANGHSISPHSEPHVLAGTVPTIGRCNISLP 1839
QY 1804 IDNGNLINVRQRKEQARGQVTVFGRRLRVHNNLSLSDFDVTRTEKIYDHRKFTLRI 1863
Db 1840 MENGINSIENRLKEQIKGVTVFGRKLRVHGRNLLSIDYDRNIRTEKIYDHRKFTLRI 1899
QY 1864 LYDQAGPSLWSBSSLNGVNVYSPGYIAGTORGIMSERMEYDOAGRITSIFADGKT 1923
Db 1900 IYDQGRPFLLWSSGLAAVNVSYFFNRLAGLQRGAMSERTIDKQRIISMFADGKV 1959
QY 1924 WSYTYLEKSVLLHSSORQYIFEPDKNDRLSSVTPMNVARQTLETIRSVGYENIYQPE 1983
Db 1960 WSYTYLEKSVLLHSSORQYIFEDSSDLHLAVTMPSVARHSMSTSVGYIENIYNPE 2019
QY 1984 GNASVIOFTEDGHLHTFVLGTGRVYIKYKLSKLAETLYDTTKVSTFYDTAGMLKT 2043
Db 2020 SNASVIFDSDGDRILKTSFGLTGRQVYKYKLSKLSLEIYDSTAVTFGYDSTGVLM 2079
QY 2044 INLQNEGFTCTIRYQIGPLIDRQIPRFTBEGMVNARFDYNY-DNSFRVTSMOAVINETP 2102
Db 2080 VNLOGSFCSTIRYKRLGPLVDKQIYRFBEGMVNARFDYTHDNFRASIKPIISETP 2139
QY 2103 LPIDLXYDVSQKTEFGKFGVYIYDINOIITAVNTHKHDPAYGRMKVQYEIFRSL 2162
Db 2140 LPVDLYRYDEISGKVEHFGFGVYIYDINOIITAVNTHKHDPAYGRMKVQYEIFRSL 2199
QY 2163 MYMTVOYDNMGRVVKELKVPYANTVTRYSYEYDADGLOQVTSINDKPLWYSYDLNGN 2222
Db 2200 MYMTVOYDNMGRVVKELKVPYANTVTRYSYEYDADGLOQVTSINDKPLWYSYDLNGN 2259
QY 2223 LHLSPGNSARLTPRYDIRDLRDLRDLRDLRDLRDLRDLRDLRDLRDLRDLRDLRDLR 2282
Db 2260 LHLNPGNSVRLMPLRYDLRDLRDLRDLRDLRDLRDLRDLRDLRDLRDLRDLRDLR 2319
QY 2283 RAGSWSRYRYDGLGRVSKSSSHHLQFVYADLTNPTKVTHLYNHSSEITSLYYDQ 2342
Db 2320 KANGWVQYRYDGLGRVSKSSSHHLQFVYADLTNPTKVTHLYNHSSEITSLYYDQ 2379
QY 2343 GHLFAMESSGDEPIYACDNIGTFLAVFSQGLMIKQILYATGEIYMDTNPNFQIIIGY 2402
Db 2380 GHLFAMESSGDEPIYACDNIGTFLAVFSQGLMIKQILYATGEIYMDTNPNFQIIIGY 2439

QY 2403 HGGLYDPLTKLVHMRDYYDLAGRWTSPDHELKHLSSNVMPPNLYMFKNNNPISNSQ 2462
Db 2440 HGGLYDPLTKLVHMRDYYDLAGRWTSPDHELKHLSSNVMPPNLYMFKNNNPISNSQ 2498
QY 2463 DIKFMPTDVNSWLLTFGQLHNVPYKPKPDMDAMFESYELHTQMTQOEWDSKSLGV 2522
Db 2499 DLKNYVTDVKSWMVFGFQSLNIIIPGFRKMYVSPPEYL-----TESQACENGQLITGV 2554
QY 2523 OCEVOKQKAFVTLERFDQLYSGTITSQQAPKTKK-----PASSGSVFGKGFALKDGR 2578
Db 2555 QOTTERHNQAFWALE-----GOVISKRLHASIREKAGHWFATSTPIIGKIMFAVKGR 2608
QY 2579 VTTDIISVANDEGRVRAAILNHAHYLENLHPTIGOVTHYFVKPSPSEGLAILGLSGGR 2638
Db 2609 VTTGSIATDSRKIASVLSNAHYLEKMHYSIEGKOTHYFVKIGSADSLSVTLAMTSGR 2668
QY 2639 RTLENGVNVTVSOINTVNLGRTRYTDIOLQYGALCNTRYG---TTLDEKARVLELAR 2695
Db 2669 KVLDSGVNVTVSQPTLLINGRTRFTNIEFOYSTLLINIRYGLTADTLDDEKARVLDQAR 2728
QY 2696 QRAVRQAWAREQORLREGSEGLRAWTEGEKQOVLSTGRVQYDGFVVISVEQYPELSDSA 2755
Db 2729 QRALGSAAWAKQOQXARDGREGSRVMTDGEKQOLLNTRGVQYEGYVYLPVEQYPELADSS 2788
QY 2756 NNHFMROSENGRR 2769
Db 2789 SNIQFLRONEMGR 2802

RESULT 15
US-10-072-012-489
; Sequence 489, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles B.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Riesger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02


```

; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 489
; LENGTH: 2802
; TYPE: PRT
; ORGANISM: Gallus gallus
US-10-072-012-489

Query Match          64.6%; Score 9616.5; DB 12; Length 2802;
Best Local Similarity 61.3%; Pred. No. 0;
Matches 1773; Conservative 411; Mismatches 491; Indels 217; Gaps 27;

QY 1 MDVKRXPYSLTARR-DAERRYTSADSEEGKAP-QKSYSSSETLKAYDQDARLAYGS 58
D 1 MDIKDRR-HRSLTEGRCCKRYTSSSLDSEDCRVPKAKSYSSSETLKAYGHDRMHYGN 59

QY 59 RVKDIVPQEAEEFCTGANETLAEGLLEEVTPPHGTLRYTDIGLPCCGYSMGAGSDADME 118
D 60 RVSOLVHRESDEFPQGTNTLAEGLGICEPS-PHRSYGCSDIGILHQGYSLTSGSDADSD 118

QY 119 ADTVLSPEHPVRLWGRTRSGRSCSLSSRANSNLTITDTEHEN-----161
D 119 TEGGMSPEHAIWLWGRGKISRSRSGLSRRENSALTITDSDENKNSDEENDFHTHLSEK 178

QY 162 -----TETH-----166
D 179 DROTSWOOLAEATKNSLIRPIPTSSSSSLPFAQLPSSHNPVPPVSCOMPLDNTSHQIM 238

QY 167 -----PGLONH-----197
D 239 DTPDEFPNSYLLRACSGPQOASSSGPSNHHQSOTLRPLPPPHNHSLS-----HHSSA 295

QY 198 NSLNRGNFTPRSN-----PSPAPTDHSLSGEPAGGAQEPAPHAQENMLLSNIPLETRNLG 253
D 296 NSLNRNSLTNRNRNQIHAPAPAPND--LATT-----ESVOLQDSWYLSNVPLETR---344

QY 254 KQPLGTQDNLIEMDLILGASRDGAYSOGHLEFK-PCGTSPLFCITSPGYPLTSTSVYS 312
D 345 -----HFUKTSSGTTPLFSSSSPGYPLTSGIVYT 374

QY 313 PPPRPLPRSTFARPAFNLKPKSKYCNWKAALSAIVISATLIVLLAYFVAMHFLGNWHL 372
D 375 PPPRLLPRNTFSRNAPFLKPKSKYCNWKAALSAIAAAVLLAILLAYFIAMHLLGLNWQL 434

QY 373 QPVEGQWYEI-----TETDASSWPVPTDVSILYPSGGTGLTDPDRKGKGTTEGKPSF 424
D 435 QPADGHTFSGNLRPGAAGEDGAAPPA-----GRG-----PW 467

QY 425 FPEDSFIDSGEIDVGRASOKIPDGTFRWRSQVFDHPVHLKFNVSIGKAALVGIYGRKGL 484
D 468 VTRNSSIDSGETEVRKVQVTEVPVQVFWRSQIHSIQOFLKFNISIGKDALFVYIRRG 527

QY 485 PPSHTQDFVELLDGRLLTOEARSLGTPRQGRGVTPPSHSTGTFIQYLDGSIWHLAFY 544
D 528 PPSHAQYDFMERLDGK-----EKWSVVESPRRRSIQTLVQNEAVFVQYLDVGLWHLAFY 582

QY 545 NDGKESEVSPLTITTAISVDNCPNSCYNGDCISGTCFGLFLGPDGCRASCPVLCSGN 604
D 583 NDGKDEKRVWFSVTILDSVQDCPNCHNGNECVSGVCHPFGPHGADCAKAAFCVLCSGN 642

QY 605 GQYMKGRCLCHSGKAGACDVPTNQICIDVACSNHGTICITGICNPGYKGCSEEVDCMD 664
D 643 GQYSKGTCLCYSGWKGPCECDVPISQCIDPSCGGHSGCIEGNCVCSIGYKGCNEEVDCLD 702

```

```

QY 665 PTCSGRGVCYVRGECHCFVWGWTNCETPRATCLDQCSGHGTFLPDTGLCSDPSWTGHC 724
D 703 PTCSNHGVCVNGECLCSPGWGGINCELPRACQPDQCSGHGTFLPDTGLCSDPNMGMGDC 762

QY 725 STEICAADCGGHGVCVGGTCEDCGWMAACDORACHPRCAEHGFCRCKGCECSGWNGE 784
D 763 SVEVCSDCGTHGVCIGGACRCEGWTGACDQRVCHPRCTEHGTCCKGCKCEGWNGE 822

QY 785 HCTIAHYLDVVVKEGCPGLCNGNGRCTLDLNGWHCVCOLGWRGAGCDTSMETACDSDKN 844
D 823 HCTIGRQITGTETDGCPLCNGNGRCTLGQNSWCQVCQGTGWRGPGCNVAMETSCADNKN 882

QY 845 DQDGLVCMDDCCLOPLCHINPLCLGSPNPLDIIQETQVPVSQONLHSPYDRIFLVR 904
D 883 EGDGLVDCLVPCCLQSTCQNSLLCRGSRDPLDIIQSSH--SGSFAVXSFYDRIKLLVVK 940

QY 905 DETHIIPGENPDDGGHACVIRGOVMTSDGTPLGVNISFVNNPLFGYITISQDGSFDLVT 964
D 941 DETHIIPGENPNSLSVSLIRGOVTTDGTPLGVNVSVFKYKYGYTITRQDGMFDLVA 1000

QY 965 NGGISIIIRFRAPFITOEHTLWLPWDRFFVNETIIMRHEENETPSCDLSNAPRNPVS 1024
D 1001 NGGSSSLTLHFRAPFMSOERTVWLPWNSFYAMDITLVKMTSENSIPSCDLSGFGVDPVII 1060

QY 1025 PSPLTSEASSCAEKPIVPEICAQEBEISISCKMRLSYLSRTPGYKSVLRISLTHPTI 1084
D 1061 SSPLTFFSDAPGRNPVIVPETQVLHEEIEVFGSSIKLIYLSRTAGYKSLKLIWQSLV 1120

QY 1085 PFNLKMHVMAVEGRLFRKFAAADLSYFIWMDTDVYNQKVFGLSEAFVSVCYEYES 1144
D 1121 PLNLKLVMAVEGHFLQKSLASPNLAYTFIWDKTDAYQKVYGLSDAVVSGFEYET 1180

QY 1145 CPDILLWEKRTVLOGYEIDASKLGGWSLDKXHALNIOGILHKGNGENQFVSQOPVIG 1204
D 1181 CPSLILWEKRTALLQGFELDFPSNLGGWSLDKXHLNWKSGILHKGNGENQLTQOPAVIT 1240

QY 1205 STMGRRRSISCPSCNGLAGDKNLLAPVALTCSGSLYVGDNFYIRRIIPSGNVNITL 1264
D 1241 STMGRRRSISCPSCNGLAEGNLLAPVALAVLGDSLFGVDNFYIRRIIPSRNVTSL 1300

QY 1265 ELNRKDPHSHSFAHXYLATDPMGSAVFLSDSNSRRVFKTKSTVVVXDLVNSVAVGT 1324
D 1301 ELNRKEFKHNNPAHXYLATDPMGSAVFLSDSNSRRVFKTKSTVVTGTLKLAGNSVAVGT 1360

QY 1325 GQOCLPFDTRCGDGGKATEATLTNPRGTVVDKGLIYFVDGTMIRRIDONGIISTLGS 1384
D 1361 GEQCLPFDTRCGDGGKATEATLTNPRGTVVDKGLIYFVDGTMIRRIDONGIISTLGS 1420

QY 1385 NDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQVIRAVGR 1444
D 1421 NDLTAVRPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQVIRAVGR 1480

QY 1445 PMHCOVPGIDHPLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIQVTTSGEIS 1504
D 1481 PMHCOVPGIDHPLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIQVTTSGEIS 1539

QY 1505 LVAGAPSGCDKNDANCDGDCGAYAKDKNTPSSLAVCADGELYVADLGNIRIRIR 1564
D 1540 LVAGAPSGCDKNDANCDGDCGAYAKDKNTPSSLAVCADGELYVADLGNIRIRIR 1599

QY 1565 KXKPFINTQNMVELSSPIDQELIYLPDITGKLYTQSLFTGDLVYNFTYTGDDITLITDN 1624
D 1600 KXKPFINTQNMVELSSPIDQELIYLPDITGKLYTQSLFTGDLVYNFTYTGDDITLITDN 1659

QY 1625 NGMNVNVRDSTGMPFLWLPVDPGVVYVMTGNTSALKSVTTQGHLEAMVTHGNSGLLAT 1684
D 1660 NGMNVNVRDSTGMPFLWLPVDPGVVYVMTGNTSALKSVTTQGHLEAMVTHGNSGLLAT 1719

QY 1685 KSNENGWTTFFBYDYSFGLTNVTPPTQGVSPRSDTSSVHVQVETSSK-DDVTTITNLS 1743
D 1720 KSNENGWTTFFBYDYSFGLTNVTPPTQGVSPRSDTSSVHVQVETSSK-DDVTTITNLS 1779

QY 1744 ASGAFYTLLOQVRSNYSYIGADGSLRLLLANGMEVALQTEPHLLAGTAVNPTVGKENVILP 1803

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 16:12:48 ; Search time 85 Seconds
(without alignments)
9204.400 Million cell updates/sec

Title: US-10-029-020-14
Perfect score: 14887
Sequence: 1 MDVKERKPYRSLTRRDABR.....ELSDSANNHFHQRSEMGSR 2769

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001s.*
5: Geneseq2002s.*
6: Geneseq2003as.*
7: Geneseq2003bs.*
8: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14887	100.0	2769	5 ABG70388	Abg70388 Human TEN
2	14748.5	99.1	2758	5 ABG97359	Abg97359 Human CGD
3	14406.5	96.8	2794	5 ABG98401	Abg98401 Human NOV
4	10221	68.7	2721	5 ABP33587	Abp33587 Human NOV
5	10201	68.5	2725	5 ABP33586	Abp33586 Human NOV
6	9888.5	66.4	2628	5 ABP33588	Abp33588 Human NOV
7	9856	66.2	2613	5 ABP33589	Abp33589 Human NOV
8	9536	64.1	2733	4 AAU08680	AAU08680 Human FCT
9	9531	64.0	2733	7 ADB32024	ADB32024 Human FCT
10	9530.5	64.0	2724	4 AAU08681	AAU08681 Human FCT
11	9530.5	64.0	2724	7 ADB32029	ADB32029 Human FCT
12	9501	63.8	2725	5 ABG61913	Abg61913 Prostate
13	9314	62.6	2633	6 ABR58318	AbR58318 BC002058
14	8874	59.6	2551	6 ABR58317	AbR58317 BC00205A
15	7296.5	49.0	2136	4 AAM78695	Aam78695 Human pro
16	6258.5	42.0	1688	6 ABR58344	AbR58344 XM_047995
17	4920	33.0	1351	6 AAO29571	Aao29571 Human Pco
18	3979.5	26.7	1045	4 AAB92858	Aab92858 Human pro
19	3806.5	25.6	2515	4 ABB71354	Abb71354 Drosophil
20	3730	25.1	964	4 AAB93294	Aab93294 Human pro
21	3302	22.2	1015	4 AAM79679	Aam79679 Human pro
22	2989	20.1	768	2 AAW44817	Aaw44817 Human gam
23	2989	20.1	768	2 AAY06639	Aay06639 Gamma-her
24	2989	20.1	768	3 AAY71176	Aay71176 Human Her
25	2989	20.1	768	4 AAU09891	Aau09891 Human her

26	2351	15.8	625	4 AAB92780	Aab92780 Human pro
27	1574	10.6	527	2 AAW44818	Aaw44818 Human gam
28	1310	8.8	402	4 AEG04674	Aeg04674 Novel hum
29	1244	8.4	459	7 ADE08090	Ado08090 Novel pro
30	1205	8.1	746	4 ABG07028	Abg07028 Novel hum
31	1184.5	8.0	777	4 ABB65598	Abb65598 Drosophil
32	1112	7.5	495	6 ADA54781	Ada54781 Human pro
33	1043	7.0	1124	4 ABB59241	Abb59241 Drosophil
34	1037.5	7.0	587	4 ABB59244	Abb59244 Drosophil
35	976	6.6	242	4 AEG04673	Aeg04673 Novel hum
36	932	6.3	429	4 AAU00392	Aau00392 Human sec
37	798	5.4	151	3 AAB40417	Aab40417 Human NOV
38	792.5	5.3	381	4 AAB61141	Aab61141 Human NOV
39	792.5	5.3	381	4 AAU08679	Aau08679 Human FCT
40	792.5	5.3	381	6 ABO23243	Abo23243 Human bre
41	792.5	5.3	381	7 ADB32022	ADB32022 Human FCT
42	682	4.6	865	4 AEG03234	Aeg03234 Novel hum
43	652	4.4	2110	7 AAO30847	Aao30847 Human cel
44	650	4.4	2108	6 ABO01379	Abo01379 Human pro
45	649.5	4.4	2167	5 ABG70361	Abg70361 Novel hum

ALIGNMENTS

RESULT 1

ABG70388
ID ABG70388 standard; protein; 2769 AA.

XX AC ABG70388;

XX DT 05-NOV-2002 (first entry)

XX DE Human TEN-M4-like protein.

XX XX

Human; NOVX; NOVX-associated disorder; cardiomyopathy; atherosclerosis; cell signal processing; metabolic pathway modulation; metabolic disorder; obesity; diabetes; infectious disease; neurodegenerative disorder; acne; Alzheimer's disease; Parkinson's disease; immune disorder; cancer; haematopoietic disorder; cirrhosis; pancreatitis; learning defect; memory defect; infertility; congenital heart defect; hair growth; pigmentation disorder; endocrine disorder; respiratory disease; health; gastro-intestinal disease; reproductive; neurological disease; bone marrow transplantation; endocrine disease; allergy; inflammation; nephrological disorder; urinary system disorder; age-related disorder; neuropsychiatric disorder; EGF-related protein; SCUBEL; TEN-M4; adipocyte complement-related C1q tumour necrosis factor; out at first; beta adrenergic receptor kinase; EphA6/ehk-2; glucose transporter; type Ia membrane suxhi-containing domain; butyrophilin; single nucleotide polymorphism.

Homo sapiens.

Key Location/Qualifiers

Misc-difference 28 /note= "Asp substituted by Gly as a result of a single

FT nucleotide polymorphism (SNP)"

Misc-difference 64

FT /note= "Val substituted by Ala as a result of a single

FT nucleotide polymorphism (SNP)"

Misc-difference 76

FT /note= "Ala substituted by Thr as a result of a single

FT nucleotide polymorphism (SNP)"

WO200257453-A2.

25-JUL-2002.

19-DEC-2001; 2001WO-US050331.

19-DEC-2000; 2000US-0256704P.

20-DEC-2000; 2000US-0257314P.

PR 02-MAY-2001; 2001US-0288153P.
 PR 29-MAY-2001; 2001US-0294075P.
 PR 24-JUL-2001; 2001US-0307506P.
 PR 10-AUG-2001; 2001US-0311590P.
 PR 10-AUG-2001; 2001US-0311613P.
 PR 29-AUG-2001; 2001US-0315617P.
 PR 14-SEP-2001; 2001US-0322335P.
 XX (CURA-) CURAGEN CORP.
 XX
 XX Gangolli EA, Patturajan M, Vernet CAM, Malyankar UM, Kekuda R;
 PI Stone DJ, Anderson D, Shimkets RA, Burgess CE, Zernhusen BD, Liu X;
 PI Spytek KA, Casman SJ, Boldog FL, Smithson G, Li L, Ji W;
 XX
 XX WPI; 2002-590744/63.
 DR N-PSDB; ABS52100.
 XX
 XX Novel isolated NOVX polypeptide useful for treating cardiomyopathy,
 PT atherosclerosis, metabolic disorders, diabetes, obesity, infectious
 PT disease, anorexia, neurodegenerative disorders, Alzheimer's disease or
 PT cancer.
 XX
 XX Claim 1; Page 53; 318pp; English.
 PS
 XX
 CC The present invention relates to new NOVX polypeptides. The invention is
 CC useful for treating or preventing a NOVX-associated disorder such as
 CC cardiomyopathy or atherosclerosis, where the disorder is related to cell
 CC signal processing and metabolic pathway modulation in a subject,
 CC preferably human. The invention is also useful for treating metabolic
 CC disorders (e.g. obesity), diabetes, infectious disease, neurodegenerative
 CC disorders (e.g. Alzheimer's disease, Parkinson's disease), immune
 CC disorders, haematopoietic disorders and various cancers. The molecules of
 CC the invention are also useful for treating or preventing cirrhosis,
 CC pancreatitis, learning and memory defects, infertility, congenital heart
 CC defects, acne, hair growth, pigmentation disorders, endocrine disorders,
 CC respiratory disease, gastro-intestinal diseases, reproductive, health,
 CC neurological diseases, bone marrow transplantation, endocrine diseases,
 CC allergy and inflammation, nephrological disorders, urinary system
 CC disorders, neuropsychiatric disorders and age-related disorders. The
 CC present amino acid sequence represents a NOVX protein of the invention
 XX
 XX Sequence 2769 AA;
 SQ
 Query Match 100.0%; Score 14887; DB 5; Length 2769;
 Best Local Similarity 100.0%; Pred No. 0;
 Matches 2769; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDVKKRPVRLRRRDAERRYSSADSEEGKAPQKYSSETLKAYDQDARLAYGRV 60
 DB 1 MDVKKRPVRLRRRDAERRYSSADSEEGKAPQKYSSETLKAYDQDARLAYGRV 60
 QY 61 KDIVPQAEBCFTGANFTLREGLVEVPPHGTLYRTDGLPQCGYSMGASDADMEAD 120
 DB 61 KDIVPQAEBCFTGANFTLREGLVEVPPHGTLYRTDGLPQCGYSMGASDADMEAD 120
 QY 121 TVLSPHPVRLWGRSTRSGSSCLSSRANSNLTLTDEHENTETDHPGGLQNHARLRTP 180
 DB 121 TVLSPHPVRLWGRSTRSGSSCLSSRANSNLTLTDEHENTETDHPGGLQNHARLRTP 180
 QY 181 PPLSHAHTPNQHAASINSINLRGNFTPRNPPAPTDHLSGEPAGGAQEPAAHQENWL 240
 DB 181 PPLSHAHTPNQHAASINSINLRGNFTPRNPPAPTDHLSGEPAGGAQEPAAHQENWL 240
 QY 241 LNSNIPIETRLNKGQFPFLQDNLTEMIDILGASRDHGDAYSOGHFLFKPGGTSPLFCTTS 300
 DB 241 LNSNIPIETRLNKGQFPFLQDNLTEMIDILGASRDHGDAYSOGHFLFKPGGTSPLFCTTS 300
 QY 301 PGYPLTSSTVSPPPRLPRSTFARFANLKPFSKYNWKAALSAIVISATLVILLAYF 360
 DB 301 PGYPLTSSTVSPPPRLPRSTFARFANLKPFSKYNWKAALSAIVISATLVILLAYF 360
 QY 361 VAMHFLGLNHLQPMQGMVEYTEDTASSWPVPTDVSPLYSGGTGLTPDRKKGKTTGK 420

DB 361 VAMHFLGLNHLQPMQGMVEYTEDTASSWPVPTDVSPLYSGGTGLTPDRKKGKTTGK 420
 QY 421 PSSFPEDSFIDSGEIDVGRRASOKIPPGTFWRSQVFDHPVHLKFNVS LGKAAVGIYG 480
 DB 421 PSSFPEDSFIDSGEIDVGRRASOKIPPGTFWRSQVFDHPVHLKFNVS LGKAAVGIYG 480
 QY 481 RKGLPSPHTQDFVELLDGRRLLTQEARSLGPTPRQSRGTVPSPSSHETGFIQYLDGSIW 540
 DB 481 RKGLPSPHTQDFVELLDGRRLLTQEARSLGPTPRQSRGTVPSPSSHETGFIQYLDGSIW 540
 QY 541 LAFYNDGKESVWSFLTTATESVDCNCPNSCYGNGDCISGTCCHQFLGFLGPDGCRASCPVL 600
 DB 541 LAFYNDGKESVWSFLTTATESVDCNCPNSCYGNGDCISGTCCHQFLGFLGPDGCRASCPVL 600
 QY 601 CSNGQYMKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTCTGTCTICNPGYKGESCEEV 660
 DB 601 CSNGQYMKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTCTGTCTICNPGYKGESCEEV 660
 QY 661 DCMPTCSGRGVCVRGCHCFVWGHTNCTPRATCLDQCSGHGTHLPDTGLCSCDPSWT 720
 DB 661 DCMPTCSGRGVCVRGCHCFVWGHTNCTPRATCLDQCSGHGTHLPDTGLCSCDPSWT 720
 QY 721 GHDCSIEICAAACGCGHGVCGGTCTCRCDGWMGAACDORACHPRCAEHGTCTCDGKCECPG 780
 DB 721 GHDCSIEICAAACGCGHGVCGGTCTCRCDGWMGAACDORACHPRCAEHGTCTCDGKCECPG 780
 QY 781 WNGEHTIAHYLDRAVVEGCGPLCNGNRCRTLDLNGWHVCQLGWRGAGCDTSMETACGD 840
 DB 781 WNGEHTIAHYLDRAVVEGCGPLCNGNRCRTLDLNGWHVCQLGWRGAGCDTSMETACGD 840
 QY 841 SKDNDGDLVDCMDPCCLQPLCHINPLCLGSPNPLDIIQETQVPVSOQNLHGFYDRIKF 900
 DB 841 SKDNDGDLVDCMDPCCLQPLCHINPLCLGSPNPLDIIQETQVPVSOQNLHGFYDRIKF 900
 QY 901 LVGRDSTHIIIPGENPFGGHACVIRGQVMTSDGTPLVGNISFVNNPLFGYTTISRDGGSF 960
 DB 901 LVGRDSTHIIIPGENPFGGHACVIRGQVMTSDGTPLVGNISFVNNPLFGYTTISRDGGSF 960
 QY 961 DLVTNGISIIILPERAPFITQETHLWLPWDRFVMEIIMRHEEHEIPSCDLSNFRPN 1020
 DB 961 DLVTNGISIIILPERAPFITQETHLWLPWDRFVMEIIMRHEEHEIPSCDLSNFRPN 1020
 QY 1021 PWSFSPSLTGFASSCAEKGPIVPEIQALQBEISISGCKMELSYLSRTPGYKSVLRISLT 1080
 DB 1021 PWSFSPSLTGFASSCAEKGPIVPEIQALQBEISISGCKMELSYLSRTPGYKSVLRISLT 1080
 QY 1081 HPTTFFNLMKVHLMVAVEGRLFRKWFAPAAPDLISYFIWKTVDVYNQKVFGLSFAFVSUGY 1140
 DB 1081 HPTTFFNLMKVHLMVAVEGRLFRKWFAPAAPDLISYFIWKTVDVYNQKVFGLSFAFVSUGY 1140
 QY 1141 EYESCPDLIIWEKRTTVLQGYEIDASKLGWSLDKHALNIQSGILHKGNGENQFVSQQP 1200
 DB 1141 EYESCPDLIIWEKRTTVLQGYEIDASKLGWSLDKHALNIQSGILHKGNGENQFVSQQP 1200
 QY 1201 PVIGSINGNRRRSISCPSCNGLADGNKLIAPVALTCGSDGSLYVGFDFNIRIFPSGNV 1260
 DB 1201 PVIGSINGNRRRSISCPSCNGLADGNKLIAPVALTCGSDGSLYVGFDFNIRIFPSGNV 1260
 QY 1261 TNLLELRNKKDFRSHSPAHKYIYATDPMSCAVFLSDSNSRVRFEKISTVTVVVDLVKNSEV 1320
 DB 1261 TNLLELRNKKDFRSHSPAHKYIYATDPMSCAVFLSDSNSRVRFEKISTVTVVVDLVKNSEV 1320
 QY 1321 VAGTGDQCLPFDDTRCGDGGKATEALTNPRTGTVDFKGLIYFVDTGTMIRRIDQNGIIST 1380
 DB 1321 VAGTGDQCLPFDDTRCGDGGKATEALTNPRTGTVDFKGLIYFVDTGTMIRRIDQNGIIST 1380
 QY 1381 LLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYLDNNVLIQISENHQVRI 1440
 DB 1381 LLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYLDNNVLIQISENHQVRI 1440
 QY 1441 VAGRPXHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIQVTTTS 1500
 DB 1441 VAGRPXHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIQVTTTS 1500

Qy 1501 GEISLVAGAPSGCDCKNDANCDCFGSDGVDGAKLNTFSSLAVCADGELYVADLGNIRI 1560
 Db 1501 GEISLVAGAPSGCDCKNDANCDCFGSDGVDGAKLNTFSSLAVCADGELYVADLGNIRI 1560
 Qy 1561 RFRKNKPELNTQMYELSPIDOEIYLFDTTCKHLYTOSLPTGVDLYNFTYTGDDITL 1620
 Db 1561 RFRKNKPELNTQMYELSPIDOEIYLFDTTCKHLYTOSLPTGVDLYNFTYTGDDITL 1620
 Qy 1621 ITDNNGMNVNRDSTGMPJLWVDPGVVYTMGTNSALKSVTTQGHELAMVTHGNSG 1680
 Db 1621 ITDNNGMNVNRDSTGMPJLWVDPGVVYTMGTNSALKSVTTQGHELAMVTHGNSG 1680
 Qy 1681 LLATKKNENGWTTFFYEDSGRLTNVTFPTGVSSFRSDTDSVHVQVETSSKDDVTITT 1740
 Db 1681 LLATKKNENGWTTFFYEDSGRLTNVTFPTGVSSFRSDTDSVHVQVETSSKDDVTITT 1740
 Qy 1741 NLSASGAFYTLLOQVNRNSYIIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGRNV 1800
 Db 1741 NLSASGAFYTLLOQVNRNSYIIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGRNV 1800
 Qy 1801 TLPIDNGLNVVWRQKEQARGQVTFGRRLRVHNRNLLSLDPRVTRTEKIYDHRKFT 1860
 Db 1801 TLPIDNGLNVVWRQKEQARGQVTFGRRLRVHNRNLLSLDPRVTRTEKIYDHRKFT 1860
 Qy 1861 LRILYDOAGRPSLWSSRLNGVNTYSPGYIAGIQRGIMSERMEYDQAGRTSRIPAD 1920
 Db 1861 LRILYDOAGRPSLWSSRLNGVNTYSPGYIAGIQRGIMSERMEYDQAGRTSRIPAD 1920
 Qy 1921 GKTWSYTYLEKSMVLLHSORQYIFEDKNDRLSSVTMPNARQTLTIRSVGYRNIYQ 1980
 Db 1921 GKTWSYTYLEKSMVLLHSORQYIFEDKNDRLSSVTMPNARQTLTIRSVGYRNIYQ 1980
 Qy 1981 PPEGNASVIOQFTEDEGHLHFTVLGTGRVVIYKXKLSKLAETLYDTTKVSPFYDETAC 2040
 Db 1981 PPEGNASVIOQFTEDEGHLHFTVLGTGRVVIYKXKLSKLAETLYDTTKVSPFYDETAC 2040
 Qy 2041 LKTNLQNEGFTCTIRYRQIGPLIDRQIFRTEEGMVNARFDYNDNSFRVTSQAVINE 2100
 Db 2041 LKTNLQNEGFTCTIRYRQIGPLIDRQIFRTEEGMVNARFDYNDNSFRVTSQAVINE 2100
 Qy 2101 TPLPIDLYRVDVSGKTEQKGKGVIIYDINQIITAVMTHTKHFAYGRMKVEQVEIFR 2160
 Db 2101 TPLPIDLYRVDVSGKTEQKGKGVIIYDINQIITAVMTHTKHFAYGRMKVEQVEIFR 2160
 Qy 2161 SLWYMTVOYDNNMGRVVKELKVGYPYANTRYSEYDADGLOTVSINDKPLWRYSDLN 2220
 Db 2161 SLWYMTVOYDNNMGRVVKELKVGYPYANTRYSEYDADGLOTVSINDKPLWRYSDLN 2220
 Qy 2221 GNHLSPGNSARLTPLRYDIRDRITRLGVDQYKMDGDFLRQGGDIFEYNSAGLLIKA 2280
 Db 2221 GNHLSPGNSARLTPLRYDIRDRITRLGVDQYKMDGDFLRQGGDIFEYNSAGLLIKA 2280
 Qy 2281 YNRAGSVRYRDGLRRVSSKSSHHLQFFYADLTNPKVTHLYNHSSEITSLIYD 2340
 Db 2281 YNRAGSVRYRDGLRRVSSKSSHHLQFFYADLTNPKVTHLYNHSSEITSLIYD 2340
 Qy 2341 LQHLFAMELSSGDEFFIACDNIGTPLAVFSGTGLMKQLIYTAIGEIVYMDTNPFIQIII 2400
 Db 2341 LQHLFAMELSSGDEFFIACDNIGTPLAVFSGTGLMKQLIYTAIGEIVYMDTNPFIQIII 2400
 Qy 2401 GYHGLYDPLTKLVHMGRRDYDLVLAGRWTSFDELHMKHLSSNMPPNLYFMKNNPISN 2460
 Db 2401 GYHGLYDPLTKLVHMGRRDYDLVLAGRWTSFDELHMKHLSSNMPPNLYFMKNNPISN 2460
 Qy 2461 SODIKCFMTDVSNLLTFGLHNVIPGYPKPDVAMEPSYELIHTCMKTQEWDNKSIL 2520
 Db 2461 SODIKCFMTDVSNLLTFGLHNVIPGYPKPDVAMEPSYELIHTCMKTQEWDNKSIL 2520
 Qy 2521 GVQCEVQKQAFVTLERFDQLXGTTTSCQAPKTKKFASGVSFGKGVKFAKDCGRVT 2580
 Db 2521 GVQCEVQKQAFVTLERFDQLXGTTTSCQAPKTKKFASGVSFGKGVKFAKDCGRVT 2580

Qy 2581 TDIISVANEDORRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGDALILGSGGRT 2640
 Db 2581 TDIISVANEDORRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGDALILGSGGRT 2640
 Qy 2641 LENGNVTVSINTVLNGRTRRYTDIOLOYGALCLNTRYGTTLDEEKARVLELARQAVR 2700
 Db 2641 LENGNVTVSINTVLNGRTRRYTDIOLOYGALCLNTRYGTTLDEEKARVLELARQAVR 2700
 Qy 2701 QAWAREOQRLREGSEGRLAWTEGEBKQVLSSTGRVQGVDFPFIISVEQYPELSDSANNIHF 2760
 Db 2701 QAWAREOQRLREGSEGRLAWTEGEBKQVLSSTGRVQGVDFPFIISVEQYPELSDSANNIHF 2760
 Qy 2761 MRQSEMGER 2769
 Db 2761 MRQSEMGER 2769

RESULT 2
 ABG97359
 ID ABG97359 standard; protein; 2758 AA.
 XX ABG97359;
 AC ABG97359;
 XX 16-DEC-2002 (first entry)
 DT 16-DEC-2002 (first entry)
 XX Human CGDD10, INCYTE 7488573CD1.
 DE Human; cell growth; differentiation; death; CGDD; cancer;
 KW Human; cell growth; differentiation; death; CGDD; cancer;
 KW cell proliferative disorder; arteriosclerosis; atherosclerosis;
 KW cirrhosis; hepatitis; paroxysmal nocturnal haemoglobinuria; psoriasis;
 KW polychaemia vera; primary thrombocytopaenia; developmental disorder;
 KW renal tubular acidosis; anaemia; mental retardation; epilepsy; AIDS;
 KW neurological disorder; Alzheimer disease; Parkinson's disease; asthma;
 KW reproductive disorder; infertility; autoimmune disorder; gout; allergy;
 KW inflammatory disorder; acquired immunodeficiency syndrome; uveitis;
 KW autoimmune thyroiditis; contact dermatitis; Crohn's disease; infection;
 KW diabetes mellitus; glomerulonephritis; irritable bowel syndrome;
 KW multiple sclerosis; osteoarthritis; osteoporosis; pancreatitis;
 KW rheumatoid arthritis.
 XX Homo sapiens.
 OS Homo sapiens.
 XX WO200272830-A2.
 FN WO200272830-A2.
 XX 19-SEP-2002.
 PD 19-SEP-2002.
 XX 08-FEB-2002; 2002WO-US003715.
 PF 08-FEB-2002; 2002WO-US003715.
 XX 09-FEB-2001; 2001US-0269111P.
 PR 23-FEB-2001; 2001US-0271175P.
 PR 08-MAR-2001; 2001US-0274503P.
 PR 09-MAR-2001; 2001US-0274552P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Yue H, Yao MG, Ison CH, Lu Y, Warren BA, Elliott VS, Baughn MR;
 XX Ding L, Xu Y, Gietzen KJ, Tang TY, Lal PG, Duggan BM, Burford N;
 PI Lu DAM, Richardson TW, Tran UK, Khare R, Walia NK;
 XX WPI: 2002-723356/78.
 DR N-PSDB; ABS78652.
 XX New human proteins associated with cell growth, differentiation and
 PT death, useful for diagnosing, treating or preventing autoimmune or
 PT inflammatory disorders (e.g. AIDS, allergy or anemia), cancer,
 PT atherosclerosis or hepatitis.
 XX Claim 1; Page 155-161; 181pp; English.
 XX The invention relates to an isolated polypeptide comprising CGDD1-12
 CC (cell growth, differentiation and death), a naturally occurring amino
 CC acid sequence at least 90% identical to CGDD, a biologically active
 CC fragment or an immunogenic fragment. Also included are the

polynucleotides encoding CGDD1-12, a recombinant polynucleotide comprising a promoter sequence operably linked to the CGDD polynucleotides, a cell transformed with the recombinant polynucleotide, a transgenic organism comprising the recombinant polynucleotide, an anti-CGDD antibody, screening for compounds which bind to/modulate or are ant/agonists of CGDD or alter the expression of CGDD polynucleotide and a CGDD polynucleotide microarray. The polypeptides, polynucleotides, agonists and antagonists are useful for diagnosing, treating or preventing disorders associated with aberrant expression of CGDD, particularly cell proliferative (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, proxymal nocturnal haemoglobinuria, polycythaemia vera, psoriasis, primary thrombocytopaenia or cancer), developmental disorders (e.g. renal tubular acidosis, anaemia or mental retardation), neurological disorders (e.g. Alzheimer disease, Parkinson's disease or epilepsy), reproductive disorders (e.g. infertility or a disruption in the menstrual cycle) or autoimmune/inflammatory disorders (e.g. AIDS, (acquired immunodeficiency syndrome) allergies, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus, glomerulonephritis, gout, irritable bowel syndrome, multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, rheumatoid arthritis, uveitis, or viral, bacterial, fungal, parasitic, protozoal or helminthic infections. They are also useful in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of proteins associated with CGDD. The present sequence represents a CGDD protein

Sequence 2758 AA:

Query Match	99.1%;	Score 14748.5;	DB 5;	Length 2758;
Best Local Similarity	99.3%;	Pred. No. 0;		
Matches 2751;	Conservative	0;	Mismatches 7;	Indels 11; Gaps 3
QY	1	MDVKERKPYRSLTRRRDAERRYTTSSADSSEGGKAPQKYSYSSSETLKAYDDADARLAYGSRV	60	
Db	1	MDVKERKPYRSLTRRRDAERRYTTSSADSSEGGKAPQKYSYSSSETLKAYDDADARLAYGSRV	60	
QY	61	KDUIVQEAEEFCRTGANFTLRELGLLEEVTPHGTLYXTDGLPQCGYSWGAGSDADMEAD	120	
Db	61	KDUIVQEAEEFCRTGANFTLRELGLLEEVTPHGTLYXTDGLPQCGYSWGAGSDADMEAD	120	
QY	121	TVLSPEHPVRLWGRSTRGRSCSLSSRANSNLITTDTEHENTETDHPGGLQNHARLRTTP	180	
Db	121	TVLSPEHPVRLWGRSTRGRSCSLSSRANSNLITTDTEHENTET - PGGQLNHARLRTTP	178	
QY	181	PPLSHAHTPNQHAASINSLRGNFTPRSNPSAPTDHSLSGBPAPAGQAEPAHAQENWL	240	
Db	179	PPLSHAHTPNQHAASINSLRGNFTPRSNPSAPTDHSLSGBPAPAGQAEPAHAQENWL	238	
QY	241	LNSNTPLETRNLKGKOFFGTQDNLNLIEMDILGASHRGAYSDDGHPFLPKCGTSPLECTTTS	300	
Db	239	LNSNTPLETRNLKGKOFFGTQDNLNLIEMDILGASHRGAYSDDGHPFLPKCGTSPLECTTTS	298	
QY	301	PGVPLTSSIVYSPPPRP;PRSTFARPAFNLKPSKYCNWKAALSAIVISATVILVLAFF	360	
Db	299	PGVPLTSSIVYSPPPRP;PRSTFARPAFNLKPSKYCNWKAALSAIVISATVILVLAFF	358	
QY	361	VAMHFLGLNWHLOPMEGWYITEDTASSVPVTDVSLYPSGGTGLTDPDRKGKGTTEK	420	
Db	359	VKGHLF - NWHLOPMEGWYITEDTASSVPVTDVSLYPSGGTGLTDPDRKGKGTTEK	416	
QY	421	PGSFPPEDSFIDSGIDVGRASOKIIPPGTFWRSOVFIHPVHLKFNWSLKGAAALGVIG	480	
Db	417	PGSFPPEDSFIDSGIDVGRASOKIIPPGTFWRSOVFIHPVHLKFNWSLKGAAALGVIG	476	
QY	481	RKGLPPSHQTQDFVSLDGRLLTQEARSLGTPRQSRGTVPSPSSHETGFIQVLDSGIWH	540	
Db	477	RKGLPPSHQTQDFVSLDGRLLTQEARSLGTPRQSRGTVPSPSSHETGFIQVLDSGIWH	536	
QY	541	LAIFYNDGKSEVWSFLTTTAIESVDNCSNCGNDCISGTCFGLFLGPDGCRASCPVL	600	
Db	537	LAFYNDGKSEVWSFLTTTAIESVDNCSNCGNDCISGTCFGLFLGPDGCRASCPVL	596	
QY	601	CSNGQYWKGRCLCHSGWGKAGCEDVPTNQCIDVACSNHGTCTTGTCICINPVGKSGCREV	660	

Db 1670 LLATKSNENGWITTFEYDSFGRLTNTVFTPTGQVSSFRSDTSSVHVQVETSSKDDVTITT 1729
 Qy 1741 NLSASGAFVTLLOQOVNSYFICADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGRNV 1800
 Db 1730 NLSASGAFVTLLOQOVNSYFICADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGRNV 1789
 Qy 1801 TLPIDNGLNLVWRQKEQARGQVTVFGRRLRVHNRNLLSLDFDRTVTRTEKIYDHRKFT 1860
 Db 1790 TLPIDNGLNLVWRQKEQARGQVTVFGRRLRVHNRNLLSLDFDRTVTRTEKIYDHRKFT 1849
 Qy 1861 LRTLYDQAGPSLWSSPSRLNGVNTYSPGYTAGIQRGIMSRMEYDQAGITSRIAD 1920
 Db 1850 LRTLYDQAGPSLWSSPSRLNGVNTYSPGYTAGIQRGIMSRMEYDQAGITSRIAD 1909
 Qy 1921 GKTWSTYILEKSMVLLHSORQVIFEDKNDRLSSVTMPNVARQTLTIRSVGYRNIYQ 1980
 Db 1910 GKTWSTYILEKSMVLLHSORQVIFEDKNDRLSSVTMPNVARQTLTIRSVGYRNIYQ 1969
 Qy 1981 PPGNASVIOQDFTEDGHLHTFYLCHGRVVIYKGLKLAETLYTKVSTYDETAM 2040
 Db 1970 PPGNASVIOQDFTEDGHLHTFYLCHGRVVIYKGLKLAETLYTKVSTYDETAM 2029
 Qy 2041 LKTIINQNEGFTCTIRYQIGPLIDROIPTFTBEGMVNARFDYVNSFRVTSMAVINE 2100
 Db 2030 LKTIINQNEGFTCTIRYQIGPLIDROIPTFTBEGMVNARFDYVNSFRVTSMAVINE 2089
 Qy 2101 TPLPIDLYRVDVSGKTEQKFGVYIYDINQIITAVMTHTKHFDAYGRMKEVQVEIIPR 2160
 Db 2090 TPLPIDLYRVDVSGKTEQKFGVYIYDINQIITAVMTHTKHFDAYGRMKEVQVEIIPR 2149
 Qy 2161 SLMYWMTVOYDNNMGRVVKELKVPYANTTRVSEYDADGQLOTVSINDKPLWRYSYDLN 2220
 Db 2150 SLMYWMTVOYDNNMGRVVKELKVPYANTTRVSEYDADGQLOTVSINDKPLWRYSYDLN 2209
 Qy 2221 GNLHLLSPGNSARLTPLRYDIRITRLGDVQYKMDDEDFLRQRGDIIFEYNSAGLLIKA 2280
 Db 2210 GNLHLLSPGNSARLTPLRYDIRITRLGDVQYKMDDEDFLRQRGDIIFEYNSAGLLIKA 2269
 Qy 2281 YNRAGSWRYRVDGLRRVSSKSSHHLQFYAOLTNPTKVTYHLNHSSEITSLYYD 2340
 Db 2270 YNRAGSWRYRVDGLRRVSSKSSHHLQFYAOLTNPTKVTYHLNHSSEITSLYYD 2329
 Qy 2341 LQHLFAMELSSGDEFYIACDNIGTPLAVPSGTLMIKQILYATGEIYMDTNPNFQII 2400
 Db 2330 LQHLFAMELSSGDEFYIACDNIGTPLAVPSGTLMIKQILYATGEIYMDTNPNFQII 2389
 Qy 2401 GYEGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHELWKLSSSNMFPNLYMFKNPNISN 2460
 Db 2390 GYEGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHELWKLSSSNMFPNLYMFKNPNISN 2449
 Qy 2461 SODIKCFMTDVNSLLTFFGLHNVIPGYPKPDMADAMESYELIHTQMTQEDNKSIL 2520
 Db 2450 SODIKCFMTDVNSLLTFFGLHNVIPGYPKPDMADAMESYELIHTQMTQEDNKSIL 2509
 Qy 2521 GVOCEVQKQKAFVTLERFDQLGSTITSCQAPKTKKFASSGSVFGKVPKALKDGRVT 2580
 Db 2510 GVOCEVQKQKAFVTLERFDQLGSTITSCQAPKTKKFASSGSVFGKVPKALKDGRVT 2569
 Qy 2581 TDIISVANEGRRVAALINHAHYLENLHFTIDGVDTHYFVKPGSPGDLAILGLSGRRT 2640
 Db 2570 TDIISVANEGRRVAALINHAHYLENLHFTIDGVDTHYFVKPGSPGDLAILGLSGRRT 2629
 Qy 2641 LENGVNVTVSQINTVLNGRTRRYTDIOLQYCALCLNTRYGTTLDEKARVLELARQAVR 2700
 Db 2630 LENGVNVTVSQINTVLNGRTRRYTDIOLQYCALCLNTRYGTTLDEKARVLELARQAVR 2689
 Qy 2701 QAWAREOQRLREGBEGLRANTEGKQOVILSTGRVQGYDGFVIVSVOYPELSDSANNIHF 2760
 Db 2690 QAWAREOQRLREGBEGLRANTEGKQOVILSTGRVQGYDGFVIVSVOYPELSDSANNIHF 2749
 Qy 2761 MQQSEMGR 2769
 Db 2750 MQQSEMGR 2758

RESULT 3
 ABB98401
 ID ABB98401 standard; protein; 2794 AA.
 XX
 AC ABB98401;
 XX
 DT 21-OCT-2002 (first entry)
 XX
 DE Human NOV1, a TEN-M4 like protein.
 XX
 KW Human; NOV1; cytostatic; Cardiant; Antiinflammatory; Immunosuppressive;
 KW Antiallergic; Haemostatic; Anti-Hiv; Antidiabetic; Anorectic;
 KW Antischismatic; Neurotropic; Hepatotropic; Neuroprotective; Nootropic;
 KW Antibacterial; Virucide; Antiparasitic; Relaxant; Anticonvulsant;
 KW Gene Therapy; NOV; cancer; heart disease; inflammation;
 KW autoimmune disorder; allergy; blood disorder; AIDS; diabetes; obesity;
 KW asthma; IGA nephropathy; cirrhosis; arthritis; Alzheimer's disease;
 KW infection; stroke; muscular dystrophy; epilepsy; wasting disorder;
 KW TEN-M4 like protein.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 379 /note= "Encoded by ATGACGATT"
 FT
 XX
 PN WO200255704-A2.
 XX
 PD 18-JUL-2002.
 XX
 PF 09-JAN-2002; 2002WO-US000554.
 XX
 PR 09-JAN-2001; 2001US-0260417P.
 PR 10-JAN-2001; 2001US-0260831P.
 PR 28-FEB-2001; 2001US-0272338P.
 PR 09-MAR-2001; 2001US-0274876P.
 PR 18-APR-2001; 2001US-0284704P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Padigar M, Li L, Zerhusen BD, Casman SJ, Shenoy S, Spytek KA;
 PI Zhong M, Gangolli BA, Burgess CE, Patturajan M, Vernet CAM;
 PI Taylor S, Tchernev VT, Miller CE, Guo X, Boldog FL, Grosse WM;
 PI Alsobrook JP, Gerlach V, Edinger S, Rothenberg ME, Ellerman K;
 PI MacDougall J, Malyankar U, Millet I, Peyman J, Smithson G;
 PI Gunther E, Stone DJ;
 XX
 DR WPI; 2002-590674/63.
 DR N-PSDB; ABB98401.
 XX
 PT NOVX polypeptides and encoding polynucleotides, useful for preventing or
 PT treating NOVX-associated disorders e.g. cancer, inflammation, or
 PT Alzheimer's disease, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS Claim 1; Page 10-11; 358pp; English.
 XX
 CC The present sequence is the protein sequence for a NOV protein. The NOV
 CC proteins and coding sequences are useful for treating or preventing NOV-
 CC associated disorders or in the manufacture of a medication for treating
 CC the disorders, such as cancer, heart disease, inflammation, autoimmune
 CC disorders, allergies, blood disorders, AIDS, diabetes, obesity, asthma,
 CC IGA nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections
 CC (e.g. bacterial, viral, parasitic), stroke, muscular dystrophy, epilepsy,
 CC and other wasting disorders associated with chronic diseases. NOV1 is a
 CC TEN-M4 like protein and the NOV1 gene is localised to chromosome 11
 XX
 SQ Sequence 2794 AA;

Query Match 96.8%; Score 14406.5; DB 5; Length 2794;
 Best Local Similarity 96.7%; Pred. No. 0;

Matches 2704; Conservative 21; Mismatches 41; Indels 31; Gaps 11;			
Qy	1	MDVKEKPYSLTRRRDAERYTSSADSEBKAPOKSYSSSETLKAYDODARLAYGSRV	60
Db	1	MDVKEKPYSLTRRRDAERYTSSADSEBKAPOKSYSSSETLKAYDODARLAYGSRV	60
Qy	61	KDIPQBAEEFCRTGANFTLREICLBEVTPHGTLYRTDGLPQCGYSGAGSDADMEAD	120
Db	61	KDIPQBAEEFCRTGANFTLREICLBEVTPHGTLYRTDGLPQCGYSGAGSDADMEAD	120
Qy	121	TVLSPEHPVRLWGRSTSRSSCLSSRANGLTDTTEHENTETDHPGGLQNHARLTPP	180
Db	121	TVLSPEHPVRLWGRSTSRSSCLSSRANGLTDTTEHENTETDHPGGLQNHARLTPP	180
Qy	181	PPLSHAHTPNQHHAAASINSNLRGNFTPRSNPSPAPTDHLSGPPAGGAQEPAAHQENWL	240
Db	181	PPLSHAHTPNQHHAAASINSNLRGNFTPRSNPSPAPTDHLSGPPAGGAQEPAAHQENWL	240
Qy	241	LNSNIPLERNLKGQPLGTLQNLNLEMDILGASRDHGDYSDGHFLFKPGGTSPLFCCTS	300
Db	241	LNSNIPLERNLKGQPLGTLQNLNLEMDILGASRDHGDYSDGHFLFKPGGTSPLFCCTS	300
Qy	301	PGYPLTSSYVSPPPRLPRSTFARPAFNLLKPSKYCNMKCAALSIVISATLIVILLAYF	360
Db	301	PGYPLTSSYVSPPPRLPRSTFARPAFNLLKPSKYCNMKCAALSIVISATLIVILLAYF	360
Qy	361	VAMHLFGLNHLQPMQMEQYEITETDASSWPVPTDVSYPGSGTGLETPDRKGKGTTECK	420
Db	361	VAMHLFGLNHLQPMQMEQYEITETDASSWPVPTDVSYPGSGTGLETPDRKGKGTTECK	420
Qy	421	PSFFPEDSFIDSGEIDVGRASQKIPPGTFMRSQVFIIDHPVHLKFNVSGLKAAALVGIYG	480
Db	421	PSFFPEASFIIDSGEIDVGRASQKIPPGTFMRSQVFIIDHPVHLKFNVSGLKAAALVGIYG	480
Qy	481	RKGLPSSHOTDFVELLDGRLLTQBARSLGTPROSGRTVPSSSHETGFIQYLDGSIWH	540
Db	481	RKGLPSSHOTDFVELLDGRLLTQBARSLGTPROSGRTVPSSSHETGFIQYLDGSIWH	540
Qy	541	LAFYNDKSEVVSFLTITAESVDNCPNCGNGDCISGTCFLGFLGPDGCRASCPLV	600
Db	541	LAFYNDKSEVVSFLTITAI--LDSW-ALCLGDGECVSGTCHCFPGFLGPDCSRACPLV	597
Qy	601	CSNGQVKGRCCLCHSGWGAECDDVTNQCIDVACSNHGTCTIGTICINPGYKGSCEB	660
Db	598	CSNGQVYKRCCLCFSGWKGTECDVPTQICDPQCGRGICIMGSCACNSYKGSCEBA	657
Qy	661	DCMDPTCSRGVCRGECFCVGMGTCNCTPRATCLDCCSGHGTPLDGTGLSCDPSWT	720
Db	658	DCIDPCCSHGVCIHGECHCSPGWGSNCBIIKTMCPDCCSGHGYLQSGSCTCDPNWT	717
Qy	721	GHDCEIICAAACGGHGVCGGTFCRCEGMAACDQACRCAEHPGTCRDGKCECSPG	780
Db	718	GPDCSNEICSDVCSGHVCGGTCRCEGWTGPACNQACRCAEHPGTCRDGKCECSPG	777
Qy	781	WNGEHTI--AHVLDRAVK--ECCPGLCNGRCCTLDLNGHVCQOLGWRGAGCDSMET	836
Db	778	WNGEHTIISLAHVLDRVWLSECCPGLCNGRCCTLDLNGHVCQOLGWRGAGCDSMET	837
Qy	837	ACGDSKNDGDGLVDCMDPCCLQPLCHINPLCLGSPNPLDIIQETQVPSQNLHSPYD	896
Db	838	ACGDSKNDGDGLVDCMDPDCC-LQPLCHINPLCLGSPNPLDIIQETQVPSQNLHSPYD	897
Qy	897	RIKFLVGRDSTHIIIPENFPDGHACVIRGQVMTSDGTPLVGNISFVNNPLFGYTIISRQ	956
Db	898	RIKFLVGRDSTHIIIPENFPDGHACVIRGQVMTSDGTPLVGNISFVNNPLFGYTIISRQ	957
Qy	957	DGSFDLVINGGSIILRFRAPFITEHTLWLPWRDFFWMEIIMRHEENEPSCDLSNF	1016
Db	958	DGSFDLVINGGSIILRFRAPFITEHTLWLPWRDFFWMEIIMRHEENEPSCDLSNF	1017
Qy	1017	ARPNPVWSPSPLTSPASSCAEKGPVPEIQALQESISISGCKOURLSYLSRTPGYKSVLR	1076
Db	1018	ARPNPVWSPSPLTSPASSCAEKGPVPEIQALQEBEISISGCKOURLSYLSRTPGYKSVLR	1077

Qy	1077	ISLTHPTTIPFNLKVLHLMVAVEGSLFRKFAAAAPDLSEYFIWTKDYVYNQKVFGLSEAFV	1136
Db	1078	ISLTHPTTIPFNLKVLHLMVAVEGSLFRKFAAAAPDLSEYFIWTKDYVYNQKVFGLSEAFV	1137
Qy	1137	SVGEYESCPDLILWEKRTTVLQGYEIDASKLGWSLDKHALNIQS-GILHKNGENQF	1195
Db	1138	SVGEYESCPDLILWEKRTTVLQGYEIDASKLGWSLDKHALNIQSGGLHKNGENQF	1197
Qy	1196	VSQPPVITGSMGNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGFDFNYIRIRF	1255
Db	1198	VSQPPVITGSMGNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGFDFNYIRIRF	1257
Qy	1256	PSGNVTNILEL--RNKOFRHSHPAHKYLLATDPMGSAVFLSDSNSRRVFKIKSTVVVKD	1313
Db	1258	PSGNVTNILELRVENKDFRHSHPAHKYLLATDPMGSAVFLSDSNSRRVFKIKSTVVVKD	1317
Qy	1314	LVKSEVVAGTGDQCLPFDDTRCGDGKGAATEALTNR--GITVDKFLIYFVDTGTMIR	1370
Db	1318	LVKSEVVAGTGDQCLPFDDTRCGDGKGAATEALTNRPGPPGIVTKFGLIYFVDTGTMIR	1377
Qy	1371	RIDQNGIISTLLGSDNLTSAAPLSCDSVMDISQVR---LEWPTDLAINPMDNSLYVLDNN	1427
Db	1378	RIDQNGIISTLLGSDNLTSAAPLSCDSVMDISQVRQVHLEWPTDLAINPMDNSLYVLDNN	1437
Qy	1428	VVLOISENHQVRIIVAGPVMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYTAETD	1487
Db	1438	VVLOISENHQVRIIVAGPVMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYTAETD	1497
Qy	1488	EKKINRIRQVTTSGEISLVAGAPSGCDCKNDANDCDFSGDDGYAKDAKLTPTSSLAVCAD	1547
Db	1498	EKKINRIRQVTTSGEISLVAGAPSGCDCKNDANDCDFSGDDGYAKDAKLTPTSSLAVCAD	1557
Qy	1548	GELYVADLGNIRIRFIRKXKPFNTQNMYSLSPIQELYLFTOTTGKHLYTQSPLTGDYL	1607
Db	1558	GELYVADLGNIRIRFIRKXKPFNTQNMYSLSPIQELYLFTOTTGKHLYTQSPLTGDYL	1617
Qy	1608	YNFYTGDDGDTLITDNNGNMNVVRDSTGMPWLVPVQGVVWMTGNTNSALKSVTTQG	1667
Db	1618	YNFYTGDDGDTLITDNNGNMNVVRDSTGMPWLVPVQGVVWMTGNTNSALKSVTTQG	1677
Qy	1668	HELAMTYHNGSGLLATKSNENGTTFEYDVSFGRLTNTVFTFGQVSSFRSDTSSVHVQ	1727
Db	1678	HELAMTYHNGSGLLATKSNENGTTFEYDVSFGRLTNTVFTFGQVSSFRSDTSSVHVQ	1737
Qy	1728	VESSKDDVTITNLSAGAFYTLQDOVNSYIIGADGSLRLLLANGMEVALQTEPHLL	1787
Db	1738	VESSKDDVTITNLSAGAFYTLQDOVNSYIIGADGSLRLLLANGMEVALQTEPHLL	1797
Qy	1788	AGTVNPTVGRNVTLPIDNGLNLVWRQRKEQARGQVTVFGRRLR--VHNRLNLLSLD	1844
Db	1798	AGTVNPTVGRNVTLPIDNGLNLVWRQRKEQARGQVTVFGRRLRVLQVHNRLNLLSLD	1857
Qy	1845	RVTREKIIDHHRKFTLILYDQAGPSLWSPSRRLNGVNVYSPGYTAGIQGIMSR	1904
Db	1858	RVTREKIIDHHRKFTLILYDQAGPSLWSPSRRLNGVNVYSPGYTAGIQGIMSR	1917
Qy	1905	MEYDQAGRTSRIFADGKTMSYTYLEK---SMVLLHSORQVIFBFDKNDRLSSVTMPN	1960
Db	1918	MEYDQAGRTSRIFADGKTMSYTYLEKAGVQSVMLLHSHQVQVIFBFDKNDRLSSVTMPN	1977
Qy	1961	VARQTLTTRSVGYRNIYQPPPEGNASVIOPTFEDGHLHLLTFLVGTGRVIVYKGLSKL	2020
Db	1978	VARQTLTTRSVGYRNIYQPPPEGNASVIOPTFEDGHLHLLTFLVGTGRVIVYKGLSKL	2037
Qy	2021	AETLYDTTKVSFTYDETAGMLKTINLQNGFTCTIYROIQGLIDRQIFRFTTEEGMVNAR	2080
Db	2038	AETLYDTTKVSFTYDETAGMLKTINLQNGFTCTIYROIQGLIDRQIFRFTTEEGMVNAR	2097
Qy	2081	FDYNDNSFRVTSQMAVINETPLPIDLYRYDDVSGKTEQFGKFGVYIYDINQIITAVMT	2140
Db	2098	FDYNDNSFRVTSQMAVINETPLPIDLYRYDDVSGKTEQFGKFGVYIYDINQIITAVMT	2157

QY	118	EADTVLSPEHPVRLWGRSTRSGRSSCLSRANSNLTLTDTHEH---TETHDPGLQNH 174
Db	121	ENEAVNSPEHANRLWGRGVKSRSSCLSRSNSALTTLTDEHNKSDGENPOQASNOQS 180
QY	175	RLRTPPPLSHAHTENOHAASINSLRNGFTPRSNPSPAPTDLHSLGSEPPAGGAQ--- 230
Db	181	TLQPLPP--SHKHOQAHH-PSITSLSNLSLTNRNQSPAP-----PAALPAELQTT 229
QY	231	-EPAHAQENWLLNSNIPLETRNLGKQPFGLTQDNLIEMDILGASRHGCAVSDGHFLPKP 289
Db	230	PESVOLQDSWLGSNVPLESR-----HFLFKT 256
QY	230	G-GTSPPLCTTSPGYPLTSSSTVSPPPPLPRSTFARPAFNLKKFSKYNMKCAALSIAIV 348
Db	257	GTGTTPLFSTATPGYTMAAGSVSPPTPLPNTILSRSAFKFKSKSKYCSWKCTALCAVG 316
QY	349	ISATLVIILAFVAVHFLGLNWHQLQPMGQMYE---ITEDASSWPVPTDVSYPGSGTG 405
Db	317	VSULLAILLSYFIAHFLGLNWLQLOQTENDTFENGKVNSDT-----MPTNVTVSLPSG--- 368
QY	406	LETDPKKGTTGTEGKPSFFPEDPSIDSGEIDVGERASQKLPBPGTFNRSOYFIDHPVHLK 465
Db	369	-----DNKLGSGFTQENTIDSGEIDIGRAIQEIPPGIFNRSQFIQPOOFLK 417
QY	466	FNVSIGKAALVGIYGRKGLPSPHTQDFVELLDGRRLLTQFARSLEGTPROSRGTVPSS 525
Db	418	FNISLQKDALIGVGRKGLPSPHTQYDFVELLDGSRLLIARQSRLLIETERAGRQARSVL 477
QY	526	HETGFIQYLDSGIWHLAFYNDGKESEVVSFLTTAIESVDNCPSCYNGDCISGTCFCFL 585
Db	478	HEAGFIQYLDSGIWHLAFYNDGKNAEQVSFNTIVIESVVECPRNCHGNCEVSGTCHCFP 537
QY	586	GFLGPDCCRASCPVLCSENGMYKGRCLCHSGWKGAECDPVTNOCIDVACSNHGTCTGT 645
Db	538	GFLGPDCSRACPVLCSENGQYSGKRCCLCFSGWKTECDVFTTCIDPQCGRGICIMGS 597
QY	646	CICNPYKGESCEBVDQMDPTCSGRGVCVRECHCFVWGNGTNCETTPATCLDCQSGHT 705
Db	598	CACSSGYKGESCEBADCIDPGCSNHGVCIIHEGCHSPGWGSGNCEILTKMCPDCQSGHT 657
QY	706	FLPDTGLCSDPSWTHGTDHCSITEICAAOCHGHGVCGVGTCTCEDGWMGAACDQRACHPRCA 765
Db	658	YLQSSGCTDPPNWTGPPDCSNEICSVDCGSHGVCMVGTCRCEBWTGPAQNQRACHPRCA 717
QY	766	EHGTCRDKGCEBSPWNGEHCIIAHYLDVVY-----KEGPGLCNGNGRCTLDLNGWHC 819
Db	718	EHGTCRDKGCEBQSGWNGEHCIIAHYLDKIVKDKIGYKEGPGLCNNGRCTLDQNGGHC 777
QY	820	VCQLGWRGAGCDTSMETACGSDKNDGDLVDCMDPCCLOPLCHINPLCLGSPNLDII 879
Db	778	VCQPGWRGACDVAMETLCTDSKDNEDGLDCMDPCCLOSSQONQPYCRGLPDPDQII 837
QY	880	QETQVPSVQNLHGFYDRIKPLVGRDSTHIIPGSNPFDGGHACVIRQVMTSDGTPUVG 939
Db	838	SQSLQSPSQAARKSFYDRIIFSLIGSDTHVIPGESPENKGLASVIRQVLTADGTPLIG 897
QY	940	NISFVNNPLPGYTISSROGSPDLVTNCGISIIILFRERAPITQBHTLWLSDWFFWMETI 999
Db	898	NVSFFHVPYGYIITROGDFDLVANGASLTLVFRSPFLTQHTVMIENWYFVMDTL 957
QY	1000	IMRHENEIPSCOLSNFARNPVVSPPLTSPASSCAEKGPVPEIQALQEEISISGCKM 1059
Db	958	VMEKEENDIPSCOLSGFVRNPVIVSSPFLSTFRSSPEDSPIPETQVLHEETIIPGTL 1017
QY	1060	RLSVLSRRTPGYKSVLRISLTHPTIPNLMKVLMVAVEGELPRKWFAPAAPDLSYFIWD 1119
Db	1018	KLVSLSRAAGYKSVLKITMQSIIIPNLMKVLMVAVGRLFQKFPFASPNIAYFIWD 1077
QY	1120	KTDVYNQKVFGLGEAFVSGYBYESCPDLILWEKRTTVLQGYEIDASKLGGWLDKXHAL 1179
Db	1078	KTDAYNOKVYGLGEAVVSGYBYESCLDLTLWEKRTALLOGYELDASNMGGWTLDKHVL 1137

Qy	1180	NIOGILHKGNGENQFVSQQPPVIGSIMGNGRRRSISCPSCNGADGNKLLAPVALTCGS	1239
Dy	1181	:::::	
Db	1138	DVOGILYKXGNGEQFTSQQPPVYSSIMGNGRRRSISCPSCNGQADGNKLLAPVALACGI	1197
Qy	1240	DGSLYVGDFNVIIRIFPSGNVTNILELRNKDPRHSHSPAHHKYLYLATDPMSGAFLVDSNS	1299
Db	1198	DGSLYVGDFNVRIIPFSGNVTSVLELRNKDPRHSSNPAAHYLYLATDPVTGDLYVSDTNT	1257
Qy	1300	RVPFKIKSTVVVKDLVKXSEVWAGTGDOCLPFDDBTRCGDGCGKATEALTINPRGITVDKFG	1359
Db	1258	RRIYRPKSLTAGAKDLTKNAEWWAGTGEQCPLPFDEARCGDGCKGAVEATLMPKGMADVDKNG	1317
Qy	1360	LIVFVDTGMIRRIDONGIISTLLGSNDLTSGARPLSCDSUMVDISOVRLEWTDLAIIMPDMN	1419
Db	1318	LIVFVDTGMIRKIDQNGIISTLLGSNDLTSGARPLTCDTSMHIISOVRLEWTDLAIIMPDMN	1377
Qy	1420	SLEYLDNNVLIQISENHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATHLESATALAVSHNG	1479
Db	1378	SIYVLDNNVLIQTENROVRIAAGRPMHCQVPGVE-YPVCKHAVQITTESATAIAVSYSG	1433
Qy	1480	VLYTAETDEKKINRIROVTTSGEISLVAGAPSCGDCKDANDANCOCFSGDDGYAKDAKLNTP	1539
Db	1437	VLYITEDEKKINRIROVTDGEISLVAGIPESTCCKNANDCOCYOSGGDYAKDAKSAP	1496
Qy	1540	SSLAVCADGELYVADLGNIREFIRKINKPFLNTQNMYELSSPIDQELYLPTDTGKHLTYQ	1599
Db	1497	SSLAASPDGLTIADLGNIKIRAVSKNKPILLNSNFVEVASPTDQELYIPDINGTHOYTV	1556
Qy	1600	SLPTGDIYNFTYTGDGOITLITDNNGMVNVRDRDSTGMPLWLVPDGVYVMTGMTNSA	1659
Db	1557	SLVTGDIYNFYSNDNDITAVTDSNGTURIIRDPRNRMPVRVUSDNQVIMULTITNGC	1616
Qy	1660	LKSVTTGCHELAMMTHGNSGLLATKSNENGWTFTEYEDSFGBLTNTVTPGTGOVSSFSD	1719
Db	1617	LKMGTAGLELVFTTHGNSGLLATKSDETGWITFPFYDSEGRLTNTVTPGTGVVTLHGHD	1676
Qy	1720	TDSVHVQVTSK-DDVTITNLASGAFYTLLOLQOVNSYYIGADGSLRLLLANGMEV	1778
Db	1677	MDKAITVDIBESSREEDVISITLSLSDISPYTWVQQLANSYQIGYDGSRLIIYASGLDS	1736
Qy	1779	ALOTEPHELLAGTVNPTVGKENVTLPIDNGLINVEMQRKEAQGOVTVGRRLRVHNRL	1838
Db	1737	HYQTEPHVLZAGTANPTVAKRNMTLPGENGNLVEMFRKXEQAQKVNVYGRKLRVNGRNL	1796
Qy	1839	LSLDFDRTVTEXIIDHRKFTLRIILYDOAGRPSLMSPSRNLGNVTVTSPGGYIAGIQR	1898
Db	1797	LSVDFDRTTKEXIIDHRKFLLRIAYDTSGHPRTLWPSSKLMVNAVITYSSTQIASIQR	1856
Qy	1999	GIMSERMEYDAQRIISRIFADOKTWSYTYLEKSMVLLLHSQSOYTFEPDKDRLSSVTM	1958
Db	1857	GTUSEKVDYDQGRIYRSRVFADOKTWSYTYLEKSMVLLLHSQSOYIFEYDMMDRLSAIM	1916
Qy	1959	PNVAROTLETIRSGVYRNIIYOPEGNASVIOQDETGHLLHTFYLGTCRRVRYKYGKULS	2018
Db	1917	PSVARBHTMQITRISIGYRNIIYNPESNASIITDYNEEGLLQTAFLCTSERVLFXYPROT	1976
Qy	2019	KLAETLYDTTKVGFYDDETAGMLKTINLQNEGFTCTIRYRQJGPLDROIFFPTEGMVN	2078
Db	1977	RLSEILIYDSTRVFYDDETAGVLKTNLQSDGFICTIRYRQJGPILDROIIFRSEDGMVN	2036
Qy	2079	ARFDYNVNSFRVTSMOAVNETPLDILYRDDVSGKTEQFGKGVIIYVDNQIITAV	2138
Db	2037	ARFDYSYDNSFRVTSMQVINETPLDILYQFDDISGKVEQFGKGVIIYDINQIISTAV	2096
Qy	2139	MTHTKHFDAYGRMKVEYOEIFRSLMYTMVTQYDNMRGVVYKELKVGYPANTTRYSEYDA	2198
Db	2097	MTYTKHFDAGRIKEIQYEIIFRSLMYWITIQYDNMRGVTKREIKIGPPANTTKYAEYDV	2156
Qy	2199	DGOLQTVSINDPLWEYSYDLNGLNHLSPGNSARLTPLRYDIRIORITPLDGVQVQMDED	2258
Db	2157	DGOLQTVYLEKIMRWNYDLNGLNHLNPSARSARLTPLYDLRORIITPLDGVQVRELDED	2216
Qy	2259	GFURQRGBDIFEYN SAGLLIKAINRAGSWSVRYRDGJGRRVSSKSSHHLQFFYADLT	2318

```

Db 2217 GFLRQGTETFEYSKGLLTRYVSKSGWTVIYRYDGLGRRVSSKTLGQHILQFFYADLT 2276
Qy 2319 NPTKVTHLVNHSSEITSLYDLOGLHLPAMELSSGDEFYIACNTGTPLAVESGTLMLK 2378
Db 2277 YPRIITHVNHSSSEITSLYDLOGLHLPAMELSSGDEFYIACNTGTPLAVESGTLMLK 2336
Qy 2379 QILYTAGYIYMDTNFQIIGYHGLGLYDPLTKLVHMGRRDYDVLVAGRWTSFDELHMKH 2438
Db 2337 QIQYTAGYIYFDSNIDFQVIGFHHGLYDPLTKLVHMGRRDYDVLVAGRWTSFDELHMKH 2396
Qy 2439 LSSNVNPNLYMFKNNPINSQDIKCPWTDVNSWLLTFGLHNVIGYKPKDMDAME 2498
Db 2397 I-KDPAFNLNFRNNPASKIHDVKYITDVSNSWLVTFGLHNAIFGFPVPFDLTE 2455
Qy 2499 PSYELIHTOMKTQEWNSKSIILGVQCEVOKLAFVTLERFQLYGSTITSQQAQPKTKK 2558
Db 2456 PSYELV-----KSQWDDIPIFGVQQQVARQAKFLSLGLGMAEVQ----VSRRRAGGAQS 2507
Qy 2559 ---FASGSVFGKVKFALKDGRVTTDIISVANEDGRRVAALINHAHYLENLHFTIDGVD 2615
Db 2508 WLFATVKSILGKVMVLSQVQVQVNLNIAEDCIKAAVNLNFAFYLENLHFTIEGKD 2567
Qy 2616 THYFVAPGPEGLAILGLSGRRILENGVNVTVSQINTVNGRTRRYTDIQLQYGALCL 2675
Db 2568 THYFIKTTPESDLTGLRLTSGKALENGINVTVSQSTTVNGRTRRFADVENQFGALAL 2627
Qy 2676 NTRYGTTLDEKARVLELAPQRAVROAMAREQORLREGEGRLPAWTEGKQQLVSTGRVQ 2735
Db 2628 HVRYGTTLDEKARVLELAPQRAVROAMAREQORLREGEGRLPAWTEGKQQLVSTGRVQ 2687
Qy 2736 GYDGFVVISVEOYPELSDSANNHFMQSMGRR 2769
Db 2688 GYDGYVVISVEOYPELSDSANNHFMQSMGRR 2721

```

RESULT 5

```

ABP33586
ID ABP33586 standard; protein; 2725 AA.
XX
AC ABP33586;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15a protein S8Q ID NO:36.
XX
KW Human; NOVX; cystostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianaemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; chromosome 4.
XX
OS Homo sapiens.
XX
FN WO200262999-A2.
XX
PD 15-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.

```

```

PR 13-APR-2001; 2001US-0283899P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malvankar U, Shimkess RA, Tchernev VT, Spaderma SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli BA, Guo X, Shenoy S;
PI Rastelli L, Swanson SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, MacDougall JR;
DR WPI; 2002-732706/79.
DR N-PSDB; ABQ82343.
XX
XX New NOVX polypeptides and polynucleotides useful for treating NOVX-
XX associated disorders, such as cancers, neurological disorders, disorders
XX of vesicular transport, gastrointestinal disorders, and autoimmune
XX diseases.
XX
XX Claim 1; Page 113; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
CC where x is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
CC cystostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianaemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC located on chromosome 4.
XX
XX Sequence 2725 AA;
XX
SQ
Query Match 68.5%; Score 10201; DB 5; Length 2725;
Best Local Similarity 66.9%; Pred No. 0;
Matches 1872; Conservative 362; Mismatches 462; Indels 102; Gaps 21;
Qy 1 MDVKERKPYRSLTR-RRDAERYTSSADSEKAP-QKSYSSSTLTKAYDD-ARLAYG 57
Db 1 MDVKERRPYCSLTGKRREKERYTNSADNEECRVPTQKSYSSSTLTKAFDSDRLLYG 60
Qy 58 SRVKDIPQEAEEFCRTGANFTLRGLGLEEVTPPHGLTVRTDGLPQCGYSMGAGSDADM 117
Db 61 NRVKDILVHREADEFTROGQNFTRLRQGVCEPATRGLAPCAEMGLPHRGYSISAGSDAT 120
Qy 118 EADTVLSPEHPVRLWGRSTRSGRSCLSRANSNLTLDTEHEN---TETDPGGLQNH 174
Db 121 ENEAVMSPEHAMELNGRGVKSGRSCLSRNSALTLDTEHENKSDSENEQPSAGSGOS 180
Qy 175 RLRTPPPLSHAHNTQHHAAINSINLRGNFTFRNPSPAPTDHLSGCEPPAGGQ---- 230
Db 181 TLQPLPP--SHKQHAQH--PSITSLNRNLSLNRNRNQSPAP-----PAALPAELQTT 229

```


Db 2337 LMLKQIQYATYGEIYFDSNIDFLVIGFEGGJYDPLTKLIHFGERDYDILAGRWTTPDIE 2396
QY 2435 LKHLSSNMPNLYMKNPNPISNSODIKCFMTDVSMLTTFQGLHNVPYGPVKPDM 2494
Db 2397 IKWRI-GKDPAPENLYMKNPNPASKIHDVKDYIIDVNSWLVTFGHLHNAIPGFPVKF 2455
QY 2495 DAMEPSYELIHTOMKTQEWNSKSIILGVCEVOKQLKAVTLERFDOLYGSTITSQQOAP 2554
Db 2456 DLTEPSYELV---KSGQWDDIPPIFGVQOQVAKAFSLGKMAEQV---VSRRRAG 2507
QY 2555 KTKK---PASSGVFCGKVKFALKDGRVTTDIISVANEDGRRAAALTNHAHYLENLHFTI 2611
Db 2508 GAOSWLFATVRLKSLGKWLVAQSVQRTVNTVNIANECIKVAALNNAFYLENLHFTI 2567
QY 2612 DGVDTYFVKPSPSEGLAILGLSGGRTLENGVNTVVSQINTVLMGRTRRTYDQLQVG 2671
Db 2568 EGRDTHYFKITTPESDLGLRLTSGKALENGINVTVSQISITVNVNGRTRRADVEMQFG 2627
QY 2672 ALCLNTRYGTLTDEEKARVLELAROAVRCAMAREOORLREGEGRATEGEKQOVLST 2731
Db 2628 ALALHVRVYGMTLDEEKARILEQARQALARAWAREOQVRDGBEGARLWTEGKROLLSA 2687
QY 2732 GRVQGVDFVIVGVEQVPELSDSANNIHFMRQSEMGR 2769
Db 2688 GKVGVDGYVLSVEQVPELADSGANNIQFLRQSEIGRR 2725
RESULT 6
ABP53588
ID ABP53588 standard; protein; 2628 AA.
AC ABP53588;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15c protein SEQ ID NO:40.
XX
KW Human; NOVX; cytotatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiarthritic; immunosuppressive; antiallergic; virucide;
KW antianemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; chromosome 4.
XX Homo sapiens.
XX
PN WO200262999-A2.
XX
PD 15-AUG-2002.
XX
XX 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 29-MAR-2001; 2001US-0279833P.
PR 13-APR-2001; 2001US-0283889P.
PR 18-APR-2001; 2001US-0284447P.
PR 25-APR-2001; 2001US-0286683P.
PR 29-MAY-2001; 2001US-0294080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 17-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX

(CURA-) CURAGEN CORP.
Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
Malyankar U, Shimkets RA, Tchernev VT, Spacerna SK, Gorman L;
Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shenoy S;
Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
Gunther E, Smithson G, Millet I, Macdougall JR;
WPI; 2002-732706/79.
DR N-PSDB; ABQ82345.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
associated disorders, such as cancers, neurological disorders, disorders
of vesicular transport, gastrointestinal disorders, and autoimmune
diseases.
PT
PS Claim 1; Page 122; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
CC antidiabetic, antiinflammatory, antirheumatic, antiarthritic, virucide,
CC immunosuppressive, antiallergic, antianemic, antibacterial, fungicide,
CC protozoacide and antihelminthic activities, and can be used in gene
CC therapy. The NOVX proteins, nucleotides or antibodies can be used in the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease selected from NOVX-associated disorder, such as cancers
CC (e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
CC ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
CC ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
CC disease), disorders of vesicular transport (e.g. cystic fibrosis,
CC diabetes mellitus, Grave's disease, or goitre), gastrointestinal
CC disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
CC autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
CC anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
CC and protozoal infections. The NOVX proteins can be used as immunogens to
CC produce antibodies and as vaccines. The NOVX nucleotide sequences may be
CC used in chromosome mapping, identifying individuals from minute
CC biological samples (tissue typing), and in forensic identification of a
CC biological sample. The present sequence represents human NOV15c, which is
CC located on chromosome 4
XX
SQ Sequence 2628 AA;
Query Match 66.4%; Score 9888.5; DB 5; Length 2628;
Best Local Similarity 65.2%; Pred. No. 0;
Matches 1821; Conservative 348; Mismatches 437; Indels 185; Gaps 21;
QY 1 MDVKERKPYRSLTR-RDAERYTSSADSEKAP-OKSYSSSETLXAYDQD-ARLAYG 57
Db 1 MDVKERRPYCSLTGRREKERYTNSADNEECRVPTOKSYSSSETLKAFDHDSSRLLYG 60
QY 58 SRVKDIPQEAEEFCRTGANFTLRELGLUEVTPPHGTLTYRTDGLPQCGYSMGAGSDADM 117
Db 61 NRVKDLVHREADEFT- 76
QY 118 EADTVLSPHPVRLWGRSTRSGRSCLSRANSNLTLDTEHTETDHPGQLQHARLR 177
Db 77 -----QECPASNOQOST-----LQ----- 90
QY 178 TPPPPLSHAHTPNQHHAAASINLNRGNFTPRNSPAPTDHSLSGEPPAGGAQ-----EP 232
Db 91 --PLPPSHKHQAQHH-PSITSLNRLNLTNRNRNQSPAP-----PAALPAELQTPPS 139
QY 233 AHAQENWLLNSNIPLETNRNLGKQFPLGLTLDQNLNLEMDILGASRHGAGSDGHFLFKPG-G 291
Db 140 VQLQDSWVLGNSVPLESR-----HFLFKGTG 166
QY 292 TSPFLCTTSPGYPLTSSVTYSPPPPLPRSTFARPAFNLKPKSKYCNWKCALSAIVISA 351
Db 167 TTFELFSTATGMSAGSVISPTPLPRNTLSRGAFAFKKSKSKYCNWKCALCAVGVSV 226
QY 352 TLVILLAYFVAMHLFGLNWLQPMQGYE---ITEDTASSWPFVPTDVSLYPSGGTGLT 408

Db 227 LLAIIISLFIANHLFGLNWQOQTENDTFENGKVSNDT-----NPTNTVSLPSG----- 275
 QY 409 PDRKGKGTTEKSPSPFPEDFIDSGETIDVGRASQKIPGOTFWRSQVFIIDHPVHLKFNV 468
 Db 276 -----DNGKLGFTQENNTIDSGELDIGRAIQEIPGIFWRSOLFIDQOFLXFN 327
 QY 469 SLGAALVGIYGRKGLPPSHTOPFVELLDGRRLLTOBASLEGTTPQSGTGVPVPSHET 528
 Db 328 SLOKDALIGVYGRKGLPPSHTOPFVELLDGRRLLTOBASLEGTTPQSGTGVPVPSHET 528
 QY 529 GFIOYLDGSIWHLAFYNDGKESVVSFLTTAIESVDNCPNCGYNGDCISGTCICLGFEL 568
 Db 388 GFIOYLDGSIWHLAFYNDGKESVVSFLTTAIESVDNCPNCGYNGDCISGTCICLGFEL 568
 QY 589 GPDGCRASCPVLCGNGQYMKGRCLCHSGWKAGBCDVPTNQCIDVACSNNHGTCTIGTCIC 648
 Db 448 GPDGCRASCPVLCGNGQYMKGRCLCHSGWKAGBCDVPTNQCIDVACSNNHGTCTIGTCIC 648
 QY 649 NPGYKGESCEBVDGMDPTCSGRGVYVGECHCFVGGTNCETPRATCLDQCCHGTFILP 708
 Db 508 NSGYKGESCEBVDGMDPTCSGRGVYVGECHCFVGGTNCETPRATCLDQCCHGTFILP 708
 QY 709 DTGLCSDDPSWTGHDCSIETCAADCGHGVYVGGTCCEDGMMGACDQACRACHPRCAEHG 768
 Db 568 ESGSCTCDPNTWGPDCSNEICSDGSHGVYVGGTCCEDGMMGACDQACRACHPRCAEHG 768
 QY 769 TCRDGKCECPGNGEHCCTIAHYLDVV-----KSGCPGLCNGNGRCTLDLNGHVCVQ 822
 Db 628 TCRDGKCECPGNGEHCCTIAHYLDVV-----KSGCPGLCNGNGRCTLDLNGHVCVQ 822
 QY 823 LGWRGAGCDVAMELTCTDSKDNEDGLIDCWDPCCLQSCCQNPFCRGLPFPDIIISOS 882
 Db 688 LGWRGAGCDVAMELTCTDSKDNEDGLIDCWDPCCLQSCCQNPFCRGLPFPDIIISOS 882
 QY 883 QVPVSQNLHSFVDRKFLVGRDSTHIIIPGPNFDDGGHCAVIRGQVMTSDGTPLVGNIS 942
 Db 748 LQSPSOQAASFYDRSFLGSDSTHIVPGESFNKSLASVIRGQVMTSDGTPLVGNIS 942
 QY 943 FVNNPFGYITISRODGSFIDVMTNGGSIILRPERAPFIQHTLWLPDRFFMETILMR 1002
 Db 808 FHYPEGYITISRODGSFIDVMTNGGSIILRPERAPFIQHTLWLPDRFFMETILMR 1002
 QY 1003 HENETIPSCDLSNAPNPVSPSPLTSFASCAEKGPIVPEIQALQEBISISGCKWRLS 1062
 Db 868 HENETIPSCDLSNAPNPVSPSPLTSFASCAEKGPIVPEIQALQEBISISGCKWRLS 1062
 QY 1063 YLSRTPGYKSLRISLTHPTIPFNLKVLHVAVEGRILFRKFAAAPDLSTYFFWDKTD 1122
 Db 928 YLSRTPGYKSLRISLTHPTIPFNLKVLHVAVEGRILFRKFAAAPDLSTYFFWDKTD 1122
 QY 1123 VYNQKVFGLSEAFVSVGYEYESCPDLILWEKRTVLQGYEIDASKLGGWSLDKHALNIQ 1182
 Db 988 VYNQKVFGLSEAFVSVGYEYESCPDLILWEKRTVLQGYEIDASKLGGWSLDKHALNIQ 1182
 QY 1183 SGLHKGNGENQVSOQPPVIGSIMNGRRRISCPSCNGLADGNKLLAPVALTCGSDGS 1242
 Db 1048 SGLHKGNGENQVSOQPPVIGSIMNGRRRISCPSCNGLADGNKLLAPVALTCGSDGS 1242
 QY 1243 LYVGDFNYIRIPPSGNVTNILELRNKPDRHSHSFAHYLLATDPMSGAVALFSDNSRRV 1302
 Db 1108 LYVGDFNYIRIPPSGNVTNILELRNKPDRHSHSFAHYLLATDPMSGAVALFSDNSRRV 1302
 QY 1303 FKIKSTVVVDLWKNSEVAGTGDQCLPDDTRCGDGGKATEATLTPRGITVDFKGLIY 1362
 Db 1168 FKIKSTVVVDLWKNSEVAGTGDQCLPDDTRCGDGGKATEATLTPRGITVDFKGLIY 1362
 QY 1363 FVDGTWIRIDONGIISTLLGNDLTSAPRLSCDSMDISQVRLBWPDTLAINPMDNSLY 1422
 Db 1228 FVDGTWIRIDONGIISTLLGNDLTSAPRLSCDSMDISQVRLBWPDTLAINPMDNSLY 1422
 QY 1423 VLDNNVVLQISNHQVRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLY 1482

Db 1288 VLDNNVVLQITENRQVRIAAGRPMHCQVPGVE-YPVGRHAVQTTLESATALAVSYGVLY 1346
 QY 1483 IAEDEKIKINRQVTTTSGEISLVAGAPSGCDKNDANCDCFSGDDGVYAKDAKNTPSL 1542
 Db 1347 IAEDEKIKINRQVTTTSGEISLVAGAPSGCDKNDANCDCFSGDDGVYAKDAKNTPSL 1542
 QY 1543 AVCADGELYVADLGNIRIRFIRKNKPPFJNTQNMVELSSPIDQELYLFDTTGKHLTYOSLP 1602
 Db 1407 AASPDTGLYIADLGNIRIRAVSKNKLNSMFEVASPTDQELYIFDINGTHQVTVSLV 1466
 QY 1603 TGDYLYNNTYTGCDGIDITLITDNGNVMNVVRSDTGMPLWLVVPGOVYVWMTGNSALKS 1662
 Db 1467 TGDYLYNNTYTGCDGIDITLITDNGNVMNVVRSDTGMPLWLVVPGOVYVWMTGNSALKS 1662
 QY 1663 VTTQCHELAMMTYHNSGLLATKSNENGTTFYBYDSFGRLTNVTFPTGOVSSFRSDTS 1722
 Db 1527 MTAQGLEVLVLYHNSGLLATKSNENGTTFYBYDSFGRLTNVTFPTGOVSSFRSDTS 1722
 QY 1723 SVHVOVETSSK-DDVTITNLSASGAFVTLLODQVRNSYIYIGADGSLRLLANGKMEVALQ 1781
 Db 1587 AITVDIESSSREEDVSITSNLSIDSFTYVQDQLRNSYQIGYDGSRLIIVASGLDSHYQ 1646
 QY 1782 TEPHLLAGTVNPTVGNKXNVTLPIDNGNLNVWRORKEOARGOVTVFGRRRLRVHNRLLSL 1841
 Db 1647 TEPHLLAGTVNPTVGNKXNVTLPIDNGNLNVWRORKEOARGOVTVFGRRRLRVHNRLLSL 1841
 QY 1842 DFDVTRTEKIYDHRKFTTILYDQAGRPSLMSPSRLNGVNVYSPGGYIAGIQRGM 1901
 Db 1707 DFDVTRTEKIYDHRKFTTILYDQAGRPSLMSPSRLNGVNVYSPGGYIAGIQRGM 1901
 QY 1902 SERMEYDQAGRTISRIPADGKTWSYTYLEKSMVLLHSORQYIFEPKNDRLSVTMPNV 1961
 Db 1767 SERMEYDQAGRTISRIPADGKTWSYTYLEKSMVLLHSORQYIFEPKNDRLSVTMPNV 1961
 QY 1962 ARQLETIRSVGYRNIYQPPENASVIOFTEDGHLHHTFYLGTRGRVYKYKSKLSLA 2021
 Db 1827 ARQLETIRSVGYRNIYQPPENASVIOFTEDGHLHHTFYLGTRGRVYKYKSKLSLA 2021
 QY 2022 ETLVDTTKVSTYDDETAGMLKTNLQNEGTCHIRIQIOPLDRIQIFRTEGMMNARF 2081
 Db 1887 ETLVDTTKVSTYDDETAGMLKTNLQNEGTCHIRIQIOPLDRIQIFRTEGMMNARF 2081
 QY 2082 DYNVDSFRVTSMAQVNETPLPIDLYRYDDVSGKTEQFGKFGVYIYDINQIITAVMTY 2141
 Db 1947 DYNVDSFRVTSMAQVNETPLPIDLYRYDDVSGKTEQFGKFGVYIYDINQIITAVMTY 2141
 QY 2142 TKHFDAYGRMKEVOYEIIFRSLMTWMTVQYDNMGRVVKELKVGYPANTTRYSEYDADGQ 2201
 Db 2007 TKHFDAYGRMKEVOYEIIFRSLMTWMTVQYDNMGRVVKELKVGYPANTTRYSEYDADGQ 2201
 QY 2202 LQTVSINDKPLWRYSDYDNLNHLSPGNSARLTPLRYDIRDRIITRIGDVOYKMDDEGFL 2261
 Db 2067 LQTVSINDKPLWRYSDYDNLNHLSPGNSARLTPLRYDIRDRIITRIGDVOYKMDDEGFL 2261
 QY 2262 RQSGDITFYNASGLIKAYNRAGSVRYRYDGLGRRVSSKSSHHLQFFVADLTNPT 2321
 Db 2127 RQSGDITFYNASGLIKAYNRAGSVRYRYDGLGRRVSSKSSHHLQFFVADLTNPT 2321
 QY 2322 KVTHLVNHSSEITSLAYDLOGLHFAWELSSGDEFYACDNTGTPLAVERSGTGLMTKQIL 2381
 Db 2187 KVTHLVNHSSEITSLAYDLOGLHFAWELSSGDEFYACDNTGTPLAVERSGTGLMTKQIL 2381
 QY 2382 YTAIGSIYMDTNPNFOIIGYHGGYLDPLTKLVHMGRRDYDVLAGRWTSPDHMLWHLSS 2441
 Db 2247 YTAIGSIYMDTNPNFOIIGYHGGYLDPLTKLVHMGRRDYDVLAGRWTSPDHMLWHLSS 2441
 QY 2442 SNVMPNLYMFKNNPINSQDIKCNFTDVNSWLLTFGQLHNVIPGYKPKDMDAMEPSY 2501
 Db 2306 SNVMPNLYMFKNNPINSQDIKCNFTDVNSWLLTFGQLHNVIPGYKPKDMDAMEPSY 2501
 QY 2502 ELIHTQMTQEWDSNKSILGVQCEVQKQKAFVTLERFDPQLYGTSITSCQQAQPKTKK--- 2558
 Db 2366 ELIHTQMTQEWDSNKSILGVQCEVQKQKAFVTLERFDPQLYGTSITSCQQAQPKTKK--- 2558
 QY 2558 ELIHTQMTQEWDSNKSILGVQCEVQKQKAFVTLERFDPQLYGTSITSCQQAQPKTKK--- 2558
 Db 2417 ELIHTQMTQEWDSNKSILGVQCEVQKQKAFVTLERFDPQLYGTSITSCQQAQPKTKK--- 2558

QY 2559 FASSGSVFGKVKALKDGRVTTDIISVANEDGRVAAIINHAHYLENLHFTIDGVDTYH 2618
Db 2418 FATVSLKLGKWLAVSGRQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQVQV 2477
QY 2619 FVKPGSPGDLAIGLSGRTPLNGVNVTVSQINTVLNGRTRRYTDIQLQYGCALNTR 2678
Db 2478 FIKTTTPESDLTGLTSLGKRALENGINVTVSQSTTVVNGRTRRFAFVDMQFQALALHVR 2537
QY 2679 VGTTLDEKARVLELARAVQAWREOQRLEGEGLRAWTEGEKQOVLSTGRVGYD 2738
Db 2538 VGMILDEKARLEQARALAWAREQQRVROGESGARLWTEGEKQQLSAGKVGQYD 2597
QY 2739 GFFVLSVQPELSDSANNHFMQSENGRR 2769
Db 2598 GYVLSVQPELSDSANNHFMQSENGRR 2628
RESULT 7
ABP53589
ID ABP53589 standard; protein; 2613 AA.
XX
AC ABP53589;
XX
DT 17-DEC-2002 (first entry)
XX
DE Human NOV15d protein SEQ ID NO:42.
XX
KW Human; NOVX; cytostatic; neuroprotective; anticonvulsant; cardiovascular;
KW cerebroprotective; nootropic; antidiabetic; antiinflammatory; fungicide;
KW antirheumatic; antiallergic; immunosuppressive; antiallergic; virucide;
KW antianemic; antibacterial; protozoacide; antihelminthic; gene therapy;
KW cancer; leukaemia; lymphoma; melanoma; neurological disorder; epilepsy;
KW stroke; ischaemic cerebrovascular disease; Alzheimer's disease; allergy;
KW Pick's disease; vesicular transport disease; cystic fibrosis; goitre;
KW diabetes mellitus; Grave's disease; gastrointestinal disorder; vaccine;
KW ulcerative colitis; gastric disorder; duodenal disorder; infection;
KW autoimmune disease; allergic reaction; autoimmune haemolytic anaemia;
KW rheumatoid arthritis; chromosome 4.
XX
OS Homo sapiens.
XX
PN WO200262999-A2.
XX
PD 15-AUG-2002.
XX
PF 31-DEC-2001; 2001WO-US049976.
XX
PR 29-DEC-2000; 2000US-0258928P.
PR 02-JAN-2001; 2001US-0259415P.
PR 04-JAN-2001; 2001US-0259785P.
PR 20-FEB-2001; 2001US-0269814P.
PR 09-MAR-2001; 2001US-0279863P.
PR 29-MAR-2001; 2001US-0279832P.
PR 13-APR-2001; 2001US-0279833P.
PR 18-APR-2001; 2001US-0283889P.
PR 25-APR-2001; 2001US-0284447P.
PR 29-MAY-2001; 2001US-0284080P.
PR 16-AUG-2001; 2001US-0312915P.
PR 17-AUG-2001; 2001US-0313325P.
PR 26-SEP-2001; 2001US-0322699P.
PR 26-NOV-2001; 2001US-0333350P.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Spytek KA, Li L, Wolenc AR, Vernet CAM, Eisen A, Liu X;
PI Malvankar U, Shinkets RA, Tchernev VT, Spaderma SK, Gorman L;
PI Kekuda R, Patturajan M, Gusev V, Gangolli EA, Guo X, Shency S;
PI Rastelli L, Casman SJ, Boldog F, Burgess CE, Edinger S, Ellerman K;
PI Gunther E, Smithson G, Millet I, Macdougall JR;
XX
DR WPI; 2002-732706/79.

DR N-PSDB; ABQ82346.
XX
PT New NOVX polypeptides and polynucleotides useful for treating NOVX-
associated disorders, such as cancers, neurological disorders, disorders
of vesicular transport, gastrointestinal disorders, and autoimmune
diseases.
XX
PS Claim 1; Page 126-127; 444pp; English.
XX
CC The present invention describes novel human proteins designated NOVX,
where X is 1 to 20 e.g. NOV1. NOVX sequences can have neuroprotective,
cytostatic, anticonvulsant, cerebroprotective, nootropic, cardiovascular,
antidiabetic, antiinflammatory, antirheumatic, antiallergic, virucide,
immunosuppressive, antihelminthic, antianemic, antibacterial, fungicide,
protozoacide and antihelminthic activities, and can be used in gene
therapy. The NOVX proteins, nucleotides or antibodies can be used in the
manufacture of a medicament for treating a syndrome associated with a
human disease selected from NOVX-associated disorder, such as cancers
(e.g. leukaemia, lymphoma, melanoma or cancer of the liver, lung, muscle,
ovary, testis and uterus), neurological disorders (e.g. epilepsy, stroke,
ischaemic cerebrovascular disease, Alzheimer's disease or Pick's
disease), disorders of vesicular transport (e.g. cystic fibrosis,
diabetes mellitus, Grave's disease, or goitre), gastrointestinal
disorders (e.g. ulcerative colitis, or gastric and duodenal disorders),
autoimmune diseases (e.g. allergic reactions, autoimmune haemolytic
anaemia, or rheumatoid arthritis), viral, bacterial, fungal, helminthic
and protozoal infections. The NOVX proteins can be used as immunogens to
produce antibodies and as vaccines. The NOVX nucleotide sequences may be
used in chromosome mapping, identifying individuals from minute
biological samples (tissue typing), and in forensic identification of a
biological sample. The present sequence represents human NOV15d, which is
located on chromosome 4
XX
SQ Sequence 2613 AA;
Query Match 66.2%; Score 9856; DB 5; Length 2613;
Best Local Similarity 65.2%; Pred. No. 0;
Matches 1815; Conservative 345; Mismatches 437; Indels 188; Gaps 21;
QY 1 MDVKERPYRLTR-RBDAERYVTSSADSEEGKAP-QKSYSSSETLKAYDQD-ARLAYG 57
Db 1 MDVKERPYRLTR-RBDAERYVTSSADSEEGKAP-QKSYSSSETLKAYDQD-ARLAYG 57
QY 58 SRVKDIPQEAEEFCRTGANFTLRELGLEEVTPPHGTYLRTDGLPQCGYMGAGSDAM 117
Db 61 NRKDLVERADEPTR----- 76
QY 118 EADTVLSPHVPVRLWGRSTRSGRSCSLSSRANSNLTLTDEHENTEDHPGGLQHARLR 177
Db 77 -----CEQPASNGQST----- 90
QY 178 TPTPLSHAHTPNQHAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGAGQ-----EP 232
Db 91 --PLPPSHKQSAQH-PSITSLNRSLTNRRNQSPAP-----PAALPAELQTTPE 139
QY 233 AHQENWLLNSNIPLETNRLGKQFFLTQDNLIEDILGASHDGNYSUGHFLFKPG-G 291
Db 140 VOLQDSWVLGNSVPLESR-----HFLFKGTG 166
QY 292 TSPPLCTTSPGYPLTSTVTSVSPPPRPLRSTFARPAFNLLKPKSKYCNWKAALSAIVISA 351
Db 167 TTEPLSATPGYTWASGVSVSPTRPLPRNTLSRSAPFKKSKSKYCSWKCTALCAVGVS 226
QY 352 TLVLLAYFVAMHLFGLNWHLOPMEGQMYE---ITEDASSWPVPTDVSIVPGSGTLET 408
Db 227 LLAILLSYFIAMHLFGLNWLQQTENDTFNGKVNSDT-----MPTNTVSLPSG----- 275
QY 409 PDRKKGKTEGKSPSPFPFDSIDVGRASQKIPFGTFRWSQVTFDHPVHLKFNV 468
Db 276 -----DNKLGFGTQENNTIDSGELDIGRAIQEIPGIFWRSOLFIDQQLFKNI 327
QY 469 SLGKAALVGIYGRKGLPPSHQTFDFVLLDORLLTOEARSLSGTPRQSGTVPSPSHET 528

Db 328 SLOKDALIGYGRKGLPSHTQYDFVELLDGSRLLIAREQORLLTERAGRAQARSUSLHEA 387
 QY 529 GFIOYLDGSIWHLAFYNDGKSEVVSFLTTAIESYDNCPSNYGNGDCISGTCCHCFLGFL 588
 Db 388 GFIOYLDGSIWHLAFYNDGKNAEQVSFNTIVIESVVECFRCHNGECVSGTCHCFGL 447
 QY 589 GPDGCRASCPLVCSGNGYMKGRCLCHSGWGACDVPTNOCIDVACSNHGTCITGTCIC 648
 Db 448 GPDGCRASCPLVCSGNGYMKGRCLCHSGWGACDVPTNOCIDVACSNHGTCITGTCIC 648
 QY 649 NPGYKGECEEEVDCMDPTCSGRGVCVRGECHCFVWGWTGNCETPRATCLDQCSGHGTFELP 708
 Db 508 NSGYKGECEEADCDIPGCSNHNVCIHGECHCFSGWGGSCNCEILKWCPCQCSGHGTYLQ 567
 QY 709 DTGLCSGDPWTHGDCSIEICAAACGGHGVGVGTCECEGWGAAACDQACHPRCAEHG 768
 Db 568 ESGSCTCDPNWGTGDCNEICVDCGSHGVGMGTCECEGWGAAACDQACHPRCAEHG 627
 QY 769 TCRGKCBGSPWNGEHCCTIAHYLDVVYKGECPGLCNGNGRCTLDLNGHVCOLGWRGA 828
 Db 628 TCRGKCBGSPWNGEHCCTIAHYLDVVYKGECPGLCNGNGRCTLDLNGHVCOLGWRGA 828
 QY 829 GCDTSMETACGDSKXNDGDLVDCMDPCCLQPLCHINPLCLGSPNPLDIIQSTQVPSQ 888
 Db 679 GCDVAMETLCTDSKXNDGDLVDCMDPCCLQSSCQNPYCRGLPDPQDIISQLOSPO 738
 QY 889 QNLHSFYDRKFLVGRDSTHIIIPGENPFGHACVIRGQVMTSDGTPFLGVNLSFVNPL 948
 Db 739 QAAKSFYDRKFLVGRDSTHIIIPGENPFGHACVIRGQVMTSDGTPFLGVNLSFVNPL 948
 QY 949 FGYSISQDSFDLVNTGGISIIILFERAPFIIQHTLWLPWDRFFVMTIIMHEBNEI 1008
 Db 799 YGVITITRQDGMFLVANGGASLTAVPERSFPLQYHTVMPVNFVYMDTLVWEKENDI 858
 QY 1009 PSCDLSNFAEPNVPVSPSPITSPASSCAEKGPIVPEIQALQBEISISGCKMRLSYLSST 1068
 Db 859 PSCDLSGFRPNPIIVSSPSTFRSPSPSPITPQVLTAEHTTIPGTDLKLSYLSST 918
 QY 1069 PGYKSVLRIISLTHPTIPFNLMKVLHVAVBGRFRKWFAPADLSYFFTWKTDVYVQV 1128
 Db 919 AGYKSVLKITMTOSIIPFNLMKVLHVAVBGRFRKWFAPADLSYFFTWKTDVYVQV 978
 QY 1129 FGLSEAVSVGYEYESCDLTLWEKRTAILQGYELDASNNGWTLDKHVLVDQNGILYK 1188
 Db 979 FGLSEAVSVGYEYESCDLTLWEKRTAILQGYELDASNNGWTLDKHVLVDQNGILYK 1038
 QY 1189 NGCENOFVSOQPPVIGSIMNGRRRISCPSCNGLADGKLLAPALVATCGSDGSIYVGD 1248
 Db 1039 NGCENOFVSOQPPVIGSIMNGRRRISCPSCNGLADGKLLAPALVATCGSDGSIYVGD 1098
 QY 1249 NYTRIRPPSNVNTLELRNKHSHSPAHKYLLATDPMGSAFLSDSNRRVFKIKST 1308
 Db 1099 NYTRIRPPSNVNTLELRNKHSHSPAHKYLLATDPMGSAFLSDSNRRVFKIKST 1158
 QY 1309 VVVKDLVKSSEVAGTDCDCLPDDTRCGGKATATLTPRGITVDFKGLIYFVDTGM 1368
 Db 1159 TGAKDLTKNAEVVAGTDCDCLPDDTRCGGKATATLTPRGITVDFKGLIYFVDTGM 1218
 QY 1369 IRRIDQNGIISTLLGNDLTSARPLSCDSWMDISQVRLEWPTDLAINPMDNSLYVLDNNV 1428
 Db 1219 IRRIDQNGIISTLLGNDLTSARPLSCDSWMDISQVRLEWPTDLAINPMDNSLYVLDNNV 1278
 QY 1429 VLOISENHQVRIVAGPMHCQVPGDHDHLLSKVATHLESATALAVSHNGVLYIAETDE 1488
 Db 1279 VLOISENHQVRIVAGPMHCQVPGDHDHLLSKVATHLESATALAVSHNGVLYIAETDE 1337
 QY 1489 KKNIRIROVTTSGEISLVAGAPSGCDKNDANDCDFSGDGDGKADKANTPSSLAACVADG 1548
 Db 1338 KKNIRIROVTTSGEISLVAGAPSGCDKNDANDCDFSGDGDGKADKANTPSSLAACVADG 1397
 QY 1549 ELYVADLGNIRIRFRKKNPFLNTQNMVELSPIDQELIYLFDTGKHLYTQSLPTGDYLY 1608
 Db 1398 TLYIADLGNIRIRAVSKKNPLNSMNFVAVSPDQELIYFDTGKHLYTQSLPTGDYLY 1457

QY 1609 NFFTYTGDGDIITLITDNGNVAVVRDSTGMPLWLVVBDGOVYVWVTMGTSALKSVTTQGH 1668
 Db 1458 NFFSYNDNDITAVTUSNGNLTIRRDPMRPMVRVSPQVQVILWTGTNGCLKGMTAQL 1517
 QY 1669 ELAMTYHGNGLLTKSNENGTTFYBYDPSFGRITNTVFTPTQGVSSFRSDTSSVHVQV 1728
 Db 1518 ELVLFYHGNGLLTKSDETCMTTFDYDSEGRITNTVFTPTQGVVNLHGMWDKAITVDI 1577
 QY 1729 ETSSK-DDVTIITNLASGAFYLLQDQVRNSYIAGDSLRLLANGMEVALQTEPHLL 1787
 Db 1578 ESSREEDSVIITSNLISSIDSFYTMVQDQRNSYIGYDGLRIIYASGLDHYQTEPHVL 1637
 QY 1788 ACTVNPYTKRNVTLPIIDNGLNLVWRQKEQARGOVTFGRRLRVHNRNLLSDFDRT 1847
 Db 1638 ACTANPTVAKRNMVTLPGENGQNLVWRFRKEQAQGVNVFGRKLRVNGRNLSSVDFRT 1697
 QY 1848 RTEKIYDHRKPTLRIYDQAGRPLMSPPSRLNGVNTYSPGYIAGIQSGIMSERMEY 1907
 Db 1698 KTEKIYDHRKPLRIAYDTSGHPTLWLPSSKLMVNVYTSSTQIASIQRTTSEKVDY 1757
 QY 1908 DOAGRITSRIFADGKTWSTYILEKSMVLLHLSQRYIFEPKNDRLSGVMTVNVARTLE 1967
 Db 1758 DOGRIVSRVADGKTWSTYILEKSMVLLHLSQRYIFEDMDRLSALTWPSVARHTMQ 1817
 QY 1968 TIRSVGYRNTIYQPEGNASVIQDFTEDGHLHTFYLGTRRRVYKYCKSLAETLYDT 2027
 Db 1818 TIRSIGYRNTIYNPESNASIITDYNBEGLLQLAFLGTSRRVFKYRRQTLSEILYDS 1877
 QY 2028 TKVSTYDGTAGMLKTNLQNEGETCTIRYRQIGPLIDRQIFRTEBGMNARPDYND 2087
 Db 1878 TRVSTYDGTAGMLKTNLQNEGETCTIRYRQIGPLIDRQIFRTEBGMNARPDYND 1937
 QY 2088 SFRVTSQAVINETPLIDLYRYDDVSKTEQFGKFGVYIYDINOIITAVMTHTKHDA 2147
 Db 1938 SFRVTSQAVINETPLIDLYRYDDVSKTEQFGKFGVYIYDINOIITAVMTHTKHDA 1997
 QY 2148 YGRKVEQYEIFRSLMYMTVQDNMGVVKKLVGPYANTTRYSEYDADQLOQTVSI 2207
 Db 1998 HGRKEQYEIFRSLMYMTVQDNMGVVKKLVGPYANTTRYSEYDADQLOQTVSI 2057
 QY 2208 NDKPLWRYSDNLGNLHLLSPGNSARLTPURYDIRDRITRLGVOYKMDDEGFLRGGD 2267
 Db 2058 NEKIMRYNYDNLGNLHLLSPGNSARLTPURYDIRDRITRLGVOYKMDDEGFLRGGD 2117
 QY 2268 IFENYAGLLIKAYNRAGSVRYRVDGLRRVSSKSHHLOFFYADLTNPKTVTHLY 2327
 Db 2118 IFENYAGLLIKAYNRAGSVRYRVDGLRRVSSKSHHLOFFYADLTNPKTVTHLY 2177
 QY 2328 NHSSSETSYLYDLOQHLFAMELSSGDEFYACDNIGTPLAVESGTGLMTKOLLYTAYGE 2387
 Db 2178 NHSSSETSYLYDLOQHLFAMELSSGDEFYACDNIGTPLAVESGTGLMTKOLLYTAYGE 2237
 QY 2388 IYMDTNPQIITGYHGGYDPLTKLHVMGRDYDVLAGEWTSPDHLMKHSNVMWPF 2447
 Db 2238 IYMDTNPQIITGYHGGYDPLTKLHVMGRDYDVLAGEWTSPDHLMKHSNVMWPF 2296
 QY 2448 NLVMEKNNPISNSQDIKCFMTDVSNSWLLTFFQLHNVI PGYPKPMDDAMESEYLIHQ 2507
 Db 2297 NLVMEKNNPISNSQDIKCFMTDVSNSWLLTFFQLHNVI PGYPKPMDDAMESEYLIHQ 2353
 QY 2508 MKTQEMDNSKILGVQCEVQKQLKAVTLERFDLYGTSITSCQAPKTKK---PASSGS 2564
 Db 2354 -KSOQWDDIPPIFVQVQVQVAKAKAFLSLGKMAEQV---VSRRAGGQASQMLWATVKS 2408
 QY 2565 VFGKGVKFKADGRVTDIISVANEDGRRAVAILNHAHYLENLHHTIDGVDFHYFKPGP 2624
 Db 2409 LICKGVNLAVSQGRVQVNTLNANEDCIKVAALNNAFYLENLHHTIEGKDTHYFIKTT 2468
 QY 2625 SEGDLAILGSGRRRLTENGVAVNTVSQINTVLNGTRRTYDIOLOYGALCLNRTGTLID 2684
 Db 2469 PESDLGTLALTSGKALENGVNTVSQITVNGTRTRPADVEMQFALALHVRIGMILD 2528

QY 2685 EEKARVLELQARAVQAWAREQORURBEGEGLRAWTEGEKQVLTSTGRVQGVDFVTS 2744
 Db EERARILEQARALARAWAREQORVRDSEGARLWTEGEKQLLSAGKVGQYGVYVLS 2588
 QY 2745 VEQPELSDSANNHFMQSEMGRR 2769
 Db 2589 VEQPELADSANNIQFURQSEIGRR 2613
 RESULT 8
 AAU08680
 ID AAU08680 standard; protein; 2733 AA.
 XX
 AC AAU08680;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human FCTR3b polypeptide sequence.
 XX
 KW Human; FCTR3; myelogenous leukaemia; carcinoma; melanoma; glioma;
 KW astrocytoma; congenital neonatal alloimmune thrombocytopenia; infection;
 KW neurological disorder; neurodegenerative disorders; nerve trauma;
 KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
 KW demyelinating Gardner syndrome; familial myelodysplastic syndrome;
 KW mental health condition; immunological disorder; allergy; infertility;
 KW bronchial asthma; Avellino type eosinophilia; lung disease; deafness;
 KW reproductive disorder; reproductive disorder; glycoprotein Ia deficiency;
 KW desmoid disease; turcot syndrome; liver cirrhosis; hepatitis C; virucide;
 KW gastric disorders; pancreatic disease; Schistosoma mansoni infection;
 KW Spinocerebellar ataxia; Plasmodium falciparum parasitaemia; diabetes;
 KW Corneal dystrophy-Greonow type I; Corneal dystrophy-lattice type I;
 KW Reis-Bucklers corneal dystrophy; cystostatic; immunosuppressive;
 KW anti-allergic; antiasthmatic; antiinfertility; antiinflammatory;
 KW antidiabetic; protozoacide; hepatotropic; virucide; ophthalmological;
 KW gynaecological; antiinfertility; immunostimulant; auditory; haemostatic;
 KW gene therapy; FCTR3b; neurestin-like protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200166747-A2.
 XX
 PD 13-SEP-2001.
 XX
 PF 05-MAR-2001; 2001WO-US007160.
 XX
 PR 03-MAR-2000; 2000US-0196592P.
 PR 03-MAR-2000; 2000US-0186718P.
 PR 06-MAR-2000; 2000US-0187293P.
 PR 06-MAR-2000; 2000US-0187294P.
 PR 17-MAR-2000; 2000US-0190400P.
 PR 07-APR-2000; 2000US-0196018P.
 PR 03-JAN-2001; 2001US-0259548P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Vernet CAM, Fernandes E, Shimkets RA, Herrmann JL, Majumder K;
 PI Macdonald J, Mishra V, Mezes PS, Rastelli L;
 XX
 DR WPI: 2001-596837/67.
 DR N-PSDB; AAS14085.
 XX
 PT Novel polypeptides designated as FCTR3 polypeptides, useful in detection,
 PT prevention and treatment of a broad range of pathological states.
 XX
 PS Claim 1; Page 35-36; 215pp; English.
 XX
 CC The invention relates to human FCTR3 polypeptides, FCTR1-FCTR7, and the
 CC nucleic acids encoding them. These sequences are useful for the treatment
 CC or prevention of numerous disorders including myelogenous leukaemia,
 CC carcinomas, melanomas, gliomas, astrocytomas, congenital neonatal
 CC alloimmune thrombocytopenia, neurological disorders, neurodegenerative
 CC disorders, nerve trauma, familial myelodysplastic syndrome, Charcot-Marie
 CC -Tooth neuropathy, demyelinating Gardner syndrome, familial

CC myelodysplastic syndrome, mental health conditions, immunological
 CC disorders, allergy and infection, bronchial asthma, Avellino type
 CC eosinophilia, lung diseases, reproductive disorders, infertility, male
 CC and female reproductive disorders, deafness, glycoprotein Ia deficiency,
 CC desmoid disease, turcot syndrome, liver cirrhosis, hepatitis C, gastric
 CC disorders, pancreatic diseases such as diabetes, Schistosoma mansoni
 CC infection, Spinocerebellar ataxia, Plasmodium falciparum parasitaemia,
 CC Corneal dystrophy-Greonow type I, Corneal dystrophy-lattice type I and
 CC Reis-Bucklers corneal dystrophy. This sequence represents FCTR3b, a
 CC neurestin-like protein
 XX
 SQ Sequence 2733 AA;
 Query Match 64.1%; Score 9536; DB 4; Length 2733;
 Best Local Similarity 61.4%; Pred. No. 0;
 Matches 1755; Conservative 414; Mismatches 476; Indels 212; Gaps 25;
 QY 1 MDVKERKPYRSLTRRR-DAERRYTSSSADSEKAP-QKSYSSSETLKYAQDQARLAYGS 58
 Db 1 MDVKRR-HRSLTRGCGKCRYTSSDSEDCTVPTQKSYSSSETLKYADHDSRMHYN 59
 QY 59 RVDIVPQEAEEFCRTGANFRLRELGLBEVTPHCTLYRTDGLPQCCYSMAGSADAME 118
 Db 60 RYTDLIHRESDFPQGTNFTLAELGICEPS-PHRSYGCDMGLHQYSJSTGSDADSD 118
 QY 119 ADTVLSPHPVRLWGRSTRSGRSSCLSRANSNLTLTDEHENTETDHPG----- 168
 Db 119 TEGGSPHARLWGRGKSRSSGLSRENSALTLTSDNENKSDDENGRPIPTSPS 178
 QY 169 -----G 169
 Db 179 LLPSAQLPSSHNPPVSCQMLLDSTSHQIMDTNPDEEFPNSYLLRACSPQOASSG 238
 QY 170 LQNH---AELRPPPLSHAHTPNOHHAASINLNRGNFTPRN---PSPAPTDSLSG 222
 Db 239 PNNHQSQTLRPLPP-PHNHTLSHH-SSANLNRSLNLTNRSQIHAPAPAND--LAT 294
 QY 223 EPPAGGAQEPHAQENWLNLSNIPLETNRLGKQPLGTLDQNLIEMDILGASRHDGAYSD 282
 Db 295 TP-----ESVQLQDSWLNLSNVPLETR----- 316
 QY 283 GHFLFK-PGTSPLFCTTSPGYPLTSSVYSPPRPLRSTPARAFNLKPKSKYCNMKC 341
 Db 317 -HFLFKTSGSTPLFSSSPGYPLTSGTYVTPPRPLRSTPARAFNLKPKSKYCNMKC 375
 QY 342 AALSAIVISATLVILLAYFVAMHFLGLNHLQPMQEGVVEITETASSWFPVDTSLXPS 401
 Db 376 AALSAIAAALLAALLAYFI-----VPSL----- 400
 QY 402 GGTGLETPDRKGGTTEGKPSFFPEDSPIDSGEIDVGRASQKIPPGTFWRSQVFIHP 461
 Db 401 -----KNSSIDSGAEVGRVTVQVPPGVFWRSHISQP 435
 QY 462 VHLKFNYSGLKAALYIGRKLPPSHTOFDVELLDGRLLTQEARSLGTPROSRGTV 521
 Db 436 QFLKFNISLUGKDALFGVYIRRLGLPPSHAQYDFMERLDGK-----EKWSVSEPRRSIQ 490
 QY 522 PPSSEHTGFIQYLDQGINHLAFYNDGKESEVVSFLTAIESVDNCPSCYNGDCISGTC 581
 Db 491 TLVQNEAVFVQYLDVGLWHLAFYNDGKEMVSFNTVLDLSVQDCPRNCHGECVSGVC 550
 QY 582 HCFGLGFLGDCGRASCPVLCGNGQYMKGRCLCHSGMKGAECDVPNQCIDVACSHNGTC 641
 Db 551 HCFPGFLGADCAKACPVLCGNGQYSGKTCQCYSGWKGAECDVPNQCIDPSCGHGSC 610
 QY 642 ITGTICINPGYKESCEEVDMDPTCSGRGVCRGECHECFVGGTTCETPTATCLDQCS 701
 Db 611 IDGNCVCSAGYKGEHEEVDCLDPTCSHGVCNGECLCSPGWGLNCELARVQCPDQCS 670
 QY 702 GHGTFPLDGLSCDPSWTHGDCSIEICAADCGGHGVCVGGTCTCEDGHWGACDORACH 761
 Db 671 GHGTYLPDTGLCSDPNMNGPDCSVEVCSVDCGTHGVCIGGACRCEEGTGAACDQRVCH 730

QY 762 PRCAHGTCTCRGKCECSGWNHGHCTTAHYLDVVVKGCGPGLCNGNGRCCTLDLNGHVCY 821
 DB 731 PRCEHGTCTCRGKCECSGWNHGHCTTAHYLDVVVKGCGPGLCNGNGRCCTLDLNGHVCY 790
 QY 822 QLVWRGACDTSMTACGDSKNDGDGLVDCWDPDCCLQPLCHNPLCLGSPNPLDIQOE 881
 DB 791 QTVWRGFCNVAMTSCADNKNEDGDLVDCDPCCLQACQNSLLCRGSRPDLIIQQ 850
 QY 882 TQVPVQONLHSPYDRKFLVGRDSTHIIPEENPFDGCHACVIRGQVMTSDGTPLVGVNI 941
 DB 851 QQ--TDMPAVKSFYDRKFLVGRDSTHIIPEENPFDGCHACVIRGQVMTSDGTPLVGVNI 908
 QY 942 SFVANNPLFGYTISSRODGSFPLVNGGSIILIRPERAFITQETHLALPWRFFVMTIIM 1001
 DB 909 SFVKPKGYTIITQDGTGFDLIANGASLTILHFERAFNSQERTVLPWMSFYAMDPLVM 968
 QY 1002 RHEENIPSCDLSNFARPNPVVSPPLTSFASCAEKGPIVPEIQALQEBISISGCKMRL 1061
 DB 969 KTEENSIPSCDLSGFRPDPPIIISPLSTFSAAPGQNPVPEIQLHEIEIPGNSVKL 1028
 QY 1062 SYLSRTPGYSVLRIISLTHPTIPNLMKVLHVAEGRLFRKWFAPAAADPLSYFFIWDKT 1121
 DB 1029 RYLSRRTAGYSKSLUKIIMTOSTVPLNIRVHLMVAVEHLFOKSFQASPLASTFIWDKT 1088
 QY 1122 DVYNQKVGLEAFVSVGYEYSCPDILWEKRTVLQGYEIDASKLGGWSLCKHAI 1181
 DB 1089 DAYQRYVGLSDAVVSVGYEYETCPSLILWEKRTALQGFELDPNSLGGWSLCKHAI 1148
 QY 1182 QSGILHKGNGENQVSPVIGSIMGNRRRSISCPSCNGLADGNKLLAPALVATCGSDG 1241
 DB 1149 KSGILHKGNGENQVSPVIGSIMGNRRRSISCPSCNGLAEGNKLAPALVAVGIDG 1208
 QY 1242 SLVYGDENYIRRRIPSGNVTNILENKPFRHSHPAKYYLATDPMGSAVFLSDNSRR 1301
 DB 1209 SLVYGDENYIRRRIPSGNVTNILENKPFRHSHPAKYYLATDPMGSAVFLSDNSRR 1268
 QY 1302 VFKISTVYVVKLVNSENSEVAGTQDCLPDDTRCGDGGKATEATITNPRGITVDFGLI 1361
 DB 1269 IYRVKLSGTGKDLAGNSEVAVGTGEQCLPFDEARCGDGGKAIDATLMSPRGIADVKNGLM 1328
 QY 1362 YFVDGTMRIRIDQNGIISTILGNDLTARPSCDSVMDISQVRLSWPTDLAINPMDNSL 1421
 DB 1329 YFVDATMIRKVDQNGIISTILGNDLTAVRPLSCDSMDVAQVRLSWPTDLAINPMDNSL 1388
 QY 1422 YVLDNNVLOI SENHVRIVAGRPVHCOVPGIDHFLLSKVAIHATLESATLAVSHNGVL 1481
 DB 1389 YVLENNVILAITENHQSIIAGRPVHCOVPGID-YSLSKLAIHSALESASATAISHTGVL 1447
 QY 1482 YIABTBKKNIRIQTTSGEISLVAGAPGCDCKNDANCDPSCDGDGAYAKALNTPSS 1541
 DB 1448 YITETDEKKNIRIQTTSGEISLVAGAPGCDCKNDANCDPSCDGDGAYAKALNTPSS 1507
 QY 1542 LAVCAGELVADLGNIRIRIFIRKKNFELTNQWVLSPIQELVFLPDTTKGHLVTSOL 1601
 DB 1508 LAVAPDGTIIADLGNIRIRIPAVSKNPLVAFNQAASPGQELVFNADGHIQTIVSL 1567
 QY 1602 PTGDYLYNFYTGDDITLITDNNGNMNVNRDSTCMPLWVPDQVYVWVTGNTSNALK 1661
 DB 1568 VTGEYLYNFYTSNDVTELDIDNNGSLKIRRDSGMPRHLLMPDNQIITLAVGTNGGLK 1627
 QY 1662 SVTTOGHELAMVTHGNSGLIATKSNENGWTFEYDPSGRITNTVTFPGQVSSFRSDTD 1721
 DB 1628 VSTQNLGLGMYTDGNTGLATKSDGTGWTTFYDHEGRITNTVTRPGVVTSLJHREWE 1687
 QY 1722 SSYHVQVETSSK-DDVTIITNLASGAFYTLQDQVRSNYSYICADGSLRLLLANGVEVAL 1780
 DB 1688 KSITIDIENSRRDDVTIITNLSSVEASTVVDQVQVRSYQLCNGNGLRVMYANGGISF 1747
 QY 1781 QTEPHLLAGTVNTVGRVNTLPIDNGLNLEWQRKEQARQVTVFGRRLVHNRNLLS 1840
 DB 1748 HSEPHVLAGTITPITIGRNLSPWENGLNLSIEWRLKEQIKGVTIFGRKLVRHGRNLLS 1807
 QY 1841 LDPDRVTRTEKIYDDHRTKFLIRILYDQAGRPVSLWSPSSRLNGVNVTVSPGGVIAGIQGI 1900

DB 1808 IDYDRNIRTEKIYDDHRTKFLIRILYDQAGRPVSLWSPSSRLNGVNVTVSPGGVIAGIQGI 1867
 QY 1901 MSERMEYDQAGRITSRIFADGKTWSYTYLLEKSMVLLHSHOQYIFEDKNDRLSVMTPN 1960
 DB 1869 MSERITDIQKGRIVSRMFADGKTWSYTYLLEKSMVLLHSHOQYIFEDKNDRLSVMTPN 1927
 QY 1961 VARQTLERISVGYVYRNIIQPPENASVIOQFTEDGHLHHTFYLGTRGRVYIKYKLSKL 2020
 DB 1928 VARHSMSTHISIGYIRNIYNPPESNASVIFDYSDGRILKTSFLGTGRQVQYKYKLSKL 1987
 QY 2021 AETLYDTTKVSTYDGTAGMLKTINLQNEGTCTIRIRIOIGLPLDRQIFRTEGMMNAR 2080
 DB 1988 SEIVYDSTAVTFGVDEITGVILKMVNLQSGGFSCTIRYRKIGLPLVDKQIYRESEEGMNA 2047
 QY 2081 FDYNY-DNSFRVTSQMAVINETPLPILYRYDDVSGKTEQKFGVYIYDINQIITAVM 2139
 DB 2048 FDYTYHDNSFRIASIKPVISETPLPVDLYRDEISGRVHEGKFGVYIYDINQIITAVM 2107
 QY 2140 THTKHFDAYGRMEVQYEIFRSLMYMTVQVNNMRVVKKELKVGPNATTRYSEYDAD 2199
 DB 2108 TLSKHFDTHGRIKVQYEMFRSLMYMTVQVNNMRVVKKELKVGPNATTRYSEYDAD 2167
 QY 2200 GLOQTVSINDPLAFRYSYDLNGLHLLSPGNSARLTPLRVDIRDRITRLGVDVQYKIDDDG 2259
 DB 2168 GLOQTVSINDPLAFRYSYDLNGLHLLSPGNSARLTPLRVDIRDRITRLGVDVQYKIDDDG 2227
 QY 2260 FLRQGGDIFFYNSAGLLIKAYNRAGSWSVRYRYDGLGRVSSKSHSHLQFFYADLTN 2319
 DB 2228 YLCQSGSDIFEYNSKGLLTRYNKAASGWSVQYRYDGVGRRASVYKTNLGHLLQFYSDLN 2287
 QY 2320 PTKVTHLYNHSSSITSLYDLOQLHAMELSSGDEYIACDNLCTPLAVPSGTGLMIKQ 2379
 DB 2288 PTRITHVYHNSBITSLYDLOQLHAMELSSGDEYIACDNLCTPLAVPSGTGLMIKQ 2347
 QY 2380 ILYTAYGEIYMDTNPNFQIIIGYHGGLYDPLTLVHMGRRDDYDLAGRWTSPDHELMKHL 2439
 DB 2348 LOYTAYGEIYDSDNPDMQWIGFPHGGLYDPLTLVHMGRRDDYDLAGRWTSPDHELMKHL 2407
 QY 2440 SSSNVMPNLYMFKNNAPIINSQDIKCMFTDVSNSWLLTFQQLHNVTIPGYPKPDMDAMEP 2499
 DB 2408 GKPEA-PFNLYMFKNNPLSELQKNYVTDVKSVMFGFQLSNIIIPGFPRAKMYFVPP 2466
 QY 2500 SYELIHTQMTQEWDNKSKILGVOCEVQKOLKAFVTLERFDQLYGSTITSQQAPKTKK- 2558
 DB 2467 PYELSEQAS-----ENQOLITGVQOTTERHNOAFMALE-----GVITTKLHASREKA 2516
 QY 2559 ---FASGSGVFGKGVKFPALDKGRVTTDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVD 2615
 DB 2517 GHWPATTPPIIGKIMFAIKRGRVTTGVSSIAEDSRKVASVLNNAVYLDKMHYSIEGKD 2576
 QY 2616 THYFVKPSPGEGDLAIIIGLGGRTLENGVNVTVSQINTVLNGRTRYTDIOI-QYGLALCL 2675
 DB 2577 THYFVK-LSADGDLVUTGTTIGRVLESVNVTVSQINTVLNGRTRYTDIOI-QYGLALCL 2636
 QY 2676 NTRYG---TTLDEBKARVLELARQVRAVQAWABEQQLREGEGELRAWTGEQQLVSTG 2732
 DB 2637 SIRYGLTPDLDEBKARVLELARQVRAVQAWABEQQLREGEGELRAWTGEQQLVSTG 2696
 QY 2733 RVQGYDGFVIVSVEQYBELSDSANNHFMQSENGR 2769
 DB 2697 RVQGYEYVLPVEQYBELADSSNTQFLRQNEGKR 2733

RESULT 9
 ADB32024
 ID ADB32024 standard; protein; 2733 AA.
 XX
 AC ADB32024;
 XX
 DT 04-DEC-2003 (first entry)
 XX
 DE Human FCTR3b polypeptide.

XX	KW	Human; FCTR3b; colorectal cancer; adenomatous polyposis coli;	QY	1	MDVKERKPYRSLTRRR-DAERRYTSSSSADSEKAP-QKSYSSSETLKAYDQDARLAYGS	58
KW	myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;	Db	1	MDVKRR-HRSLTRGRCKEYKTSSSLDSEDCRVPTQKSYSSSETLKAYDSDRSMHYGN	59	
KW	ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;	QY	59	RVKDIVPOAEAEFCRTGANFTLRGLLEBVTTPHGTLYRTDGLPQCGYSMGAGSDADME	118	
KW	renal cell carcinoma; melanoma; clear cell carcinoma;	Db	60	RVTDLIHREDEFPQGTNFTLAEIGICEPS-PHRSGYCSDMGLHQYSLTSGDADSD	118	
KW	granular cell carcinoma; neurological disorder;	QY	119	ADTVLSPHPVRLMGRSTRSGRSSCLSRANSNLTLTTEHENTETDHPG-----	168	
KW	neurodegenerative disorder; nerve trauma;	Db	119	TEGMSPEHAIRLMGRGKSRSSGLSRNSALTLTSDNENKSDDENGRPIPTSSPS	178	
KW	familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;	QY	169	-----G	169	
KW	Gardner syndrome; mental health condition; immunological disorder;	Db	179	LLPSAQLPSSHNPVSCVCOMPLDSDNTSHQIMDTPDEEFPNSVLLRACSGPQQAASSG	238	
KW	allergy; asthma; lung disease; reproductive disorder; deafness;	QY	170	LQNH---ARLRTPPPLSHAHTPNQHHAAASINLNRGNFTPSN-----PSPAPTDSHLSG	222	
KW	glycoprotein deficiency; desmoid tumour; turcot syndrome;	Db	239	PNHHSQSLRPLPP-PHNHTLSHHH--SSANSLNRNSLTNRRSQIHAPAPAPND--LAT	294	
KW	liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;	QY	223	EPAGGAQEPAAQENWLLNSNIPLETENLCKQPLGLTQDNLIEMDILGASRHDGAYSD	282	
KW	diabetes; schistosoma mansoni infection; spinocerebellar ataxia;	Db	295	TP-----ESVOLQDSWLVNSNPVLETR-----	316	
KW	plasmodium falciparum infection; Groenouw's corneal dystrophy;	QY	283	GHELFK-PGTSPLFCTTSPGYPLTSTVYSPPRPLPRSTFARPAFNLKKPSKYNWKC	341	
KW	lattice corneal dystrophy.	Db	317	-HELKFTSSGSTPLFSSSSPGYPLTSGTVYTPPELLPRNTEFSAKPAFLKKPSKYCSMKC	375	
XX	OS	Homo sapiens.	QY	342	AALSIVISATLVILLAYFVAMHLFGLNWHLOPMEGQMYETEDTASSWPVPTDVSUPLS	401
XX	XX	US2003087816-A1.	Db	376	AALSAIAAALLAILLAYFI-----VPMSL-----	400
XX	XX	08-MAY-2003.	QY	402	GGTLETDPDRKGGTTEGKSPFPEDSFIDSGEIDVGRASQKIPPGTFRWSQVFI DHP	461
XX	XX	05-MAR-2001; 2001US-00800198.	Db	401	-----KNSIDSGAEVGRVTVQEVPPGVFWSQJHIQSP	435
XX	XX	03-MAR-2000; 2000US-0186592P.	QY	462	VHLKPNVSLGKAALVGIYGRKGLPSPHTQDFVELLDGRRLLTQBARSLGTEPPROSRTV	521
XX	PA	(VERM/) VERMET C.	Db	436	QFLKFNISLIGKDALFGVYIRGLPSPHAQYDFMERLDGK-----EKWSVVEFRRRIQ	490
XX	PA	(FERN/) FERNANDES E.	QY	522	PPSHTGTQYLDSDGIWHLAFYNDGKESEVVSFLTITTAIESVDNCPSNCGYNGDCISGTC	581
XX	PA	(SHIM/) SHIMKETS R.	Db	491	TLVQNEAVFQYLDVGLMHLAFYNDGKEMVSNFVTVLDSVQDCPRNCHNGECVSGVC	550
XX	PA	(HERR/) HERRMANN J.	QY	582	HCFLGLGPDGCRASCPVLCSGNGQYMKGRCLCHSGWKGAECDFVPTNCIDVACSNHGTG	641
XX	PA	(MAJU/) MAJUMDER K.	Db	551	HCFFPGLGADCAKAAACPVLCSGNGQYSGTQCQYSGWKGAECDFVPMNCIDPSCGCHGSC	610
XX	PA	(MACD/) MACDOUGALL J.	QY	642	ITGTCICNPGYKGESCEVEVDCMDPTCSGRGVVGRGECHCFVGGWGTNCETPTATCLDQCS	701
XX	PA	(MISH/) MISHRA V.	Db	611	IDGNCVCSAGYKGEHCEVEVDCMDPTCSSHGVVNGECLCSPGWGLNCELARVQCPDQCS	670
XX	PA	(MEZE/) MEZES P S.	QY	702	GHGTFLPDTGLCSGDPMSWTGHDCEIPICAADCGHGVVGGTCRCEDGWMGAACDQACH	761
XX	PA	(ZAST/) RASTELLI L.	Db	671	GHGTFLPDTGLCSGDPMSWTGHDCEIPICAADCGHGVVGGTCRCEDGWMGAACDQACH	730
XX	PI	Vermet C, Fernandes E, Shinkets R, Herrmann J, Majumder K;	QY	762	PCRAHGTGCRDCKCSCPGWNGEHTCIAHLDVVVKEGCGPLCNGNCRCTLDLNHWCVC	821
XX	PI	Maddougall J, Mishra V, Mezes PS, Rastelli L;	Db	731	PCRIEHTGCKDKGCKCREGNGEHTCIQRTAGTETDGPDLCLNGNCRCTLQNSWQVCV	790
XX	XX	WPI; 2003-625633/59.	QY	822	QLGWRGAGCDTSMETACGSKNDGDLVDCMDPCCLQPLCHINPLCLGSPNPDIIQE	881
XX	XX	N-PSDB; ADB32023.	Db	791	QWGWGPGCNVAMETSCADNKNEDGVLVDCMDPCCLQSGNLLCRGRDPLDIIQQ	850
XX	PT	New FCTR polypeptide and encoding polynucleotide, useful for preventing	QY	882	TQVPSQONLHSFYDRIRKFLVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLVGVNI	941
XX	PT	or treating FCTR-related disorders, such as cancer, autoimmune,	Db	851	GQ--TDWPAVKSFYDRIRKLLAGKDSHTIIIPGENPNSLSVLRQVVTDTGTPLVGVNV	908
XX	PT	neurodegenerative, gastrointestinal, reproductive and inflammatory	QY	942	SVNNPLFGYITISRODGSFDLVNTNGGISIIIRFERAPFITQEHITLWLPDNDFFVNETIIM	1001
XX	PS	diseases.	Db	909	SPVKPKYGTITRODGTDFLIANGGASLTLLHFERAPFMSQERTVWLDPWNSFYAMDTLVM	968
XX	PS	Claim 1; Page 32-33; 155pp; English.				
XX	CC	The invention relates to FCTR polypeptides and the polynucleotides				
XX	CC	encoding them. The sequences of the invention are useful for the				
XX	CC	manufacture of a medicament for diagnosing and treating disorders				
XX	CC	associated with the FCTR polypeptide, such as colorectal cancer,				
XX	CC	adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal				
XX	CC	autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain				
XX	CC	tumours, mammary tumours, human gliomas, astrocytomas, renal cell				
XX	CC	carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear				
XX	CC	cell and granular cell carcinomas, neurological disorders,				
XX	CC	neurodegenerative disorders, nerve trauma, familial myelodysplastic				
XX	CC	syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health				
XX	CC	conditions, immunological disorders, allergy and infection, asthma, lung				
XX	CC	diseases, male and female reproductive disorders, deafness, glycoprotein				
XX	CC	deficiency, desmoid tumour, turcot syndrome, liver cirrhosis, hepatitis				
XX	CC	C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni				
XX	CC	infection, spinocerebellar ataxia, plasmodium falciparum infection,				
XX	CC	Groenouw's corneal dystrophy and lattice corneal dystrophy. This sequence				
XX	CC	represents an FCTR polypeptide of the invention.				
XX	XX	Sequence 2733 AA;				
XX	XX	Query Match 64.0%; Score 9531; DB 7; Length 2733;				
XX	XX	Best Local Similarity 61.4%; Pred. No. 0;				
XX	XX	Matches 1755; Conservative 413; Mismatches 477; Indels 212; Gaps 25;				

Db 1200 SLVGDENYIRIPPSRNVTSILELRNKEFKSNNPAHKYLLAVDPVSGSLYVSDTNSRR 1259
QY 1302 VFELKSTVVKVLDVKNSEVAGTGDOCLPDDTRCGDGGKATEATLTNPRGITVDKFLG 1361
Db 1260 IYRVKLSUGTKDLAGNSEVAGTGEQCLPFDEARCGDGGKADATDILMSPRGIAVDKNGLM 1319
QY 1362 YFVDGTMIRRIDQNGIISTLLGNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSL 1421
Db 1320 YFVDATMIRKVDQNGIISTLLGNDLTSARPLSCDSMDVAQVRLEWPTDLAINPMDNSL 1379
QY 1422 YVLDNNVQLISENHQVRIAGRPMPHQVPGIDHFLLSKVAIHATLESATALAVSHGVL 1481
Db 1380 YVLENNVILRITENHQVSIITAGRPMPHQVPGID-YLSKLAIHSALESASATISHTGVL 1438
QY 1482 YIABTEKKINRIRQVTTSGEISLVAGAPSGCCKNDANDCFSGDDGVAKAKLNTSPS 1541
Db 1439 YIETDEKKINRIRQVTTNGEICLLAGAASDCCKNDVNCVSGDDAYATDAILNPS 1498
QY 1542 LAVCAGELVADLGNIRIRIRKXPNFLNTQMYELSSPIDOBYLFTDTGKHLXTQSL 1601
Db 1499 LAVAPDGTIYIADLGNIRIRAVSKNFPVLNAFNQYEAASPEGEOLYVFNADGHIQVTVSL 1558
QY 1602 PTGDLNFTYTGDDITLLTDNNGNMVNRDSTGMLVLPVDPGVVYVMTGMTNSALK 1661
Db 1559 VTGEYLNFTYSTDNDVTEILDNNGSLKIRRSSGMPRHLLMPDQIITLIVTINGGLK 1618
QY 1662 SVTQGHLEAMTYHNGSGLLTKSNENGMWTFTEYDPSFRLTNVFTPTGVSSPSRSDTD 1721
Db 1619 VVSTQNLGLMTYDNGTGLLTKSDBTGWTFTEYDHEGRLTNVTRPTGVVTSLHREME 1678
QY 1722 SSVHVQVETSSK-DDVTITNTNLASGAFYTLLODOVRNSYIAGDSIRLLLANGNEVAL 1780
Db 1679 KSTIDIENSRDDVTIITNLSSEVASYTVVODQVRNSYQLCNCNLTFLVYANGGISF 1738
QY 1781 QTEPHLAGVNTVPTVGRNVTLPIDNGLNLEVRQKQARGOVTVFGRRLVNRNLLS 1840
Db 1739 HSEPHVLAGTITPTIGRCNISLPWENGLNIEWRLKEQIKGKVTIFGRKLAVHGNLLS 1798
QY 1841 LDFRVTREKIIDHRKFTLRILYDQAGPPLSWSPSRRLNGVNVYSPGGVIAGIQRGI 1900
Db 1799 IDYDRNIRTEKIIDHRKFTLRILYDQVGRPPLWLPSSGLAAVNVSYFFNGSLAGLQGA 1858
QY 1901 MSERMEYDQAGRTSIRFADGKTWSYTYLEKSVLLLSORQYIIPFDKNDLSVVTMEN 1960
Db 1859 MSERTDIDKGRIVSRMFADGKTWSYLYDKSVLLLSQORQYIIPYDSDDELLAVTMS 1918
QY 1961 VARTQLETIRSVGYRNIIYOPPEGNASVIOFPTDGHLLHTFVLGTGRVRIYKYKLSKL 2020
Db 1919 VASHMSHTSISGIRNIYNPPESNASVIFDYSDDGRILKTSFLGTGRQVYKYKLSKL 1978
QY 2021 AETLYDPTKVSFYDETAGMLKTNILQNEGFTCTIRYRIGPLIDRQIERFTEEGVNRAR 2080
Db 1979 SEIVDSTAVTFGDETTGVLRKWNVLQSGFSCTIRYRIGPLVDKQIRYFEEBGNVNR 2038
QY 2081 FDYNY-DNSPRVTSMAVINETPLIDLXYRDDVSGKTEQFGKFGVIYYDINQIITAVM 2139
Db 2039 FDVYTHDNSPRIASIKPVISETPLPVDLYRYDEISCKVEHFGKFGVIYYDINQIITAVM 2098
QY 2140 THYKHDAVGRMEKVOYEIFRSLMYMTVQYDNGRUVVKELKVGPNYANTTYSVEYDAD 2199
Db 2099 TLSKGFHDTGRIKEVQYEMFRSLMYMTVQYDSMGVIRKELKVGPNYANTTITYDYDGD 2158
QY 2200 GOLQTVSINDKPLWRYSYDLNGLNHLSPCNSARLTPLRYDIRDRITRLGDDVQYXMDGD 2259
Db 2159 GOLQSVAVNDRPTWRYSYDLNGLNHLNPGNSVRLMPLAYDLDRITRLGDDVQYXIDDDG 2218
QY 2260 FLQORGDIPEYNSAGLLIKAYNRAGSVRYRYDGLGRVRSKSHHLOFFVYADLTN 2319
Db 2219 YLCQSGSDIPEYNSKGLLTRYAVKASGWSVQYRYDGVGRGRASYKTNLGHLLQYFYSDLHN 2278
QY 2320 PTKVTHLYNHSSEITSYLYDLOCHLFAMELSSGDEFFIACDNIGTPTPLAVFGTGLMLIKQ 2379
Db 2279 PTKVTHLYNHSSEITSYLYDLOCHLFAMESSGGEYVVASDNTGTPLAVFINGMLIKQ 2338

QY 2380 ILTYAYGEIYMDTNPFOIIGYHGGLYDPLTKLVHMGRRDYLVLGRWTSPPDHELWKHL 2439
Db 2339 LOYTAYGEIYDSDNPQWVIGFHGGLYDPLTKLVHFTQRDYDVLGRWTSPPDYTMKNV 2398
QY 2440 SSSNVMPENLYMFKNNPINSODIKCFMTDVNSWLLTFQQLHNVITPGYKPKPDMADAVEP 2499
Db 2399 GKPEA-PENLYMFKNNPINSSELDLKNVTDVKSVMFQFQLSNIIIPGFPRAKMYVPP 2457
QY 2500 SYELIHKMKTQEWDNKSIILGVQCEVQKQKQKAPVTLERFDOLYGSTITSQQQAPKTK- 2558
Db 2458 PYESESQAS-----ENGQITGVQQTTERHNOAFMALE-----GOVITKKLHASIREKA 2507
QY 2559 ---PASSGSVFGKGVKFAKLDGRVTTDIIISVANEEDGRRVAAILNHAHYLENLHFTIDGVD 2615
Db 2508 GHWFATTTPIIGKIMFAIKREGRVTTGCVSSIASDSRKSASVLNNAAYLDMHYSIEGKD 2567
QY 2616 THYFVKGPSPSEGLAAILGLSGRRTLENGVNVTVSQINTVLNGRTRYTOIQOYGMALCL 2675
Db 2568 THYFVKIGSADGDLVTTGTTIGRKVLESQVNVTVSQPTLLVNGRTRYTOIEFYSTLL 2627
QY 2676 NTRYG---TTLDEKARVLELAROAVRQAWAREQQRLREGEGLRAWTGEQKQVLSTG 2732
Db 2628 SIRGLTPDITLDEKARVLDQARQALGTAWAKEQKARDGREGSRLWTEGEKQQLSTG 2687
QY 2733 RVQYDGFVVISVQYPELSDSANNIHFMRQSENGRR 2769
Db 2688 RVQYEGYVLPVQYPELADSSNIOFLRQENMGKR 2724
RESULT 11
ADB32029
ID ADB32029 standard; protein; 2724 AA.
XX
AC ADB32029;
DT 04-DEC-2003 (first entry)
DE Human FCTR3f polypeptide.
XX
KW Human; FCTR3f; colorectal cancer; adenomatous polyposis coli;
KW myelogenous leukaemia; congenital neonatal autoimmune thrombocytopenia;
KW ovarian cancer; brain tumour; breast cancer; glioma; astrocytoma;
KW renal cell carcinoma; melanoma; clear cell carcinoma;
KW granular cell carcinoma; neurological disorder;
KW neurodegenerative disorder; nerve trauma;
KW familial myelodysplastic syndrome; Charcot-Marie-Tooth neuropathy;
KW Gardner syndrome; mental health condition; immunological disorder;
KW allergy; asthma; lung disease; reproductive disorder; deafness;
KW glycoprotein deficiency; desmoid tumour; turcot syndrome;
KW liver cirrhosis; hepatitis C; gastric disorder; pancreatic disease;
KW diabetes; schistosoma mansoni infection; spinocerebellar ataxia;
KW plasmodium falciparum infection; Groenouw's corneal dystrophy;
KW lattice corneal dystrophy.
XX
OS Homo sapiens.
XX
FN US2003087816-A1.
XX
PD 08-MAY-2003.
XX
PF 05-MAR-2001; 2001US-00800198.
XX
PR 03-MAR-2000; 2000US-0186592P.
XX
PA (VERM/) VERMET C.
PA (FERN/) FERNANDES E.
PA (SHIM/) SHIMKETS R.
PA (HERR/) HERRMANN J.
PA (MAJU/) MAJUMDER K.
PA (MACD/) MACDOUGALL J.
PA (MISH/) MISHRA V.
PA (MEZE/) MEZES P S.

(RAST//) RASTELLI L.
 Vermet C, Fernandes E, Shinkets R, Herrmann J, Majumder K;
 Macdougall J, Mishra V, Mezes PS, Rastelli L;
 WPI; 2003-625633/59.
 N-PSDB; ADB32028.
 New FCTR polypeptide and encoding polynucleotide, useful for preventing
 or treating FCTR-related disorders, such as cancer, autoimmune,
 neurodegenerative, gastrointestinal, reproductive and inflammatory
 diseases.
 Claim 1; Page 37-38; 155pp; English.
 The invention relates to FCTR polypeptides and the polynucleotides
 encoding them. The sequences of the invention are useful for the
 manufacture of a medicament for diagnosing and treating disorders
 associated with the FCTR polypeptide, such as colorectal cancer,
 adenomatous polyposis coli, myelogenous leukaemia, congenital neonatal
 autoimmune thrombocytopenia, malignant ovarian tumours, malignant brain
 tumours, mammary tumours, human gliomas, astrocytomas, renal cell
 carcinoma, breast adenocarcinoma, ovarian carcinoma, melanomas, clear
 cell and granular cell carcinomas, neurological disorders,
 neurodegenerative disorders, nerve trauma, familial myelodysplastic
 syndrome, Charcot-Marie-Tooth neuropathy, Gardner syndrome, mental health
 conditions, immunological disorders, allergy and infection, asthma, lung
 diseases, male and female reproductive disorders, deafness, glycoprotein
 deficiency, desmoid tumour, teratoid syndrome, liver cirrhosis, hepatitis
 C, gastric disorders, pancreatic diseases, diabetes, schistosoma mansoni
 infection, spinocerebellar ataxia, plasmodium falciparum infection,
 Grounau's corneal dystrophy and lattice corneal dystrophy. This sequence
 represents an FCTR polypeptide of the invention.
 Sequence 2724 AA;

Query Match 64.0%; Score 9530.5; DB 7; Length 2724;
 Best Local Similarity 61.4%; Pred. No. 0;
 Matches 1755; Conservative 414; Mismatches 467; Indels 221; Gaps 26;
 1 MDVKERYPLSLTERR--DAERRYSSADSEEGKAP--QKSYSSSETLKAYDQDARLAYGS 58
 1 MDVKERR--HSLTRGCKGECRYTSSLDSEDCVPTQKSYSSSETLKAYDHSRMYGN 59
 59 RVKDIVPQEAEEFCORTGANTRELGLIEVTPPHGLTYRTDGLPQCGYMGAGSDADME 118
 60 RVTDLIHRESDEFPQGTNFTLAELGICEPS--PHRSGYCSDMGILHQGYSLSTGSDADS 118
 119 ADTVLSEHPVRLWGRSTRSGSSCLSSRANSNLTLDTEHENTETDHPG----- 168
 119 TEGMSPEHAIRLWGRIGKRRSGLSRENSALTLDSDNENKSDDENGRPIPTSSPS 178
 169 -----G 169
 179 LLPSAQLPSSHNPVPPVSCQPLLDNSNTSHQIMDNPDEEFPNSYLLRACSGPQQASSG 238
 170 LQNH---ARLTPPPPSHAHTPNQHAASINLRGNFTPSN----PSAPTDSLSG 222
 239 PPNHSSQSTLRPLPP--PHNHTLSHHH--SSANSNRNSLTNRRSQIHAPAPND--LAT 294
 223 EPPAGGAQEPAAHQENMLNSNIPLETNRNLGKQPFLGTLDNLIEMDILGASRHDGAYS 282
 295 TP-----ESVQLQDSVNLNSNVLETR----- 316
 283 GHFLFK--PGGTSPLCTTSBGYPYTSSTVYSSPPRPLRSTFTFAPAPNLKKPKSYCNWKC 341
 317 -HFLFKTSSTSTPLFSSSSPGYPLTSGTYTPPPRLLPRNTFSRKAFKFKPKSYCSWKC 375
 342 AALSALVISAFLVILLAYFVAMHLFGLNWHLPQMEGQYVEITDASSWFPVPTDVSLYPS 401
 376 AALSALAAALLAILAIYFI-----VFWSL----- 400
 402 GGTGLETPDRKGKGTGKPSFFPEDSFIDSGBDVGRRASQKIPPGTFWRSQVFDHP 461

Db 401 -----KNSISDSGEAEVGRVTVQVPPGVFWRSQHIHQP 435
 QY 462 VHLKFNVLGKAALVGIYGRKGLPSPHTQDFVELLDQRRLLTQBARLEGTFPRSGTV 521
 Db 436 QFLKENISLGDALFQVIRRGLPSPHAQYDFMERLDGK-----EKWSVSPRERRSTQ 490
 QY 522 PSSHSHETGFIQVLDGSIWHLAFYNDGKESVVSFLTTAIESVNDGSPSCYNGDCISGTC 581
 Db 491 TLUVQNEAFVQVLDVGLWHLAFYNDGKEMVSNFTVLDVSQDCPRCHNGGECVSGVC 550
 QY 582 HCFGLFLGPDGGRASCFLVLCGNGQYMRGRLCHSGWKGABCDVPTNCIDVACSNHGTC 641
 Db 551 HCFPGFLGADCAKAAACPVLCGNGQYSGTQCYSGWKGABCDVPMNQIDPSCGSHGSC 610
 QY 642 ITGTCINPCKYKESCEEVDCMDPTCSGRGVCVRGSECHCFVGWGTNCTETPATCLDQCS 701
 Db 611 IDGNCVCSAGYKGECEEVDCMDPTCSSHGVGVNGECLCSPGWGLNCELARVQCPDQCS 670
 QY 702 GHTGTLPTDGLSCDPPSWTGHDCSTIEICAADCGGHGVGVGTCRCEDEGMMGAACQORACH 761
 Db 671 GHTGTLPTDGLSCDPPNMWGPDCSVGVCSVDCGTHGVGIGGACRCEEGWTAACDQVCH 730
 QY 762 PRCAEHGTCRDCKSCSPGWNGEHTIAHYLDVVVKEGCPGLCNGNGRCTLDLNWHVC 821
 Db 731 PRCEIHGTCCKGCKCREGNGEHTI-----DGPDLNCGNGRCTLQONSWQVC 781
 QY 822 QLGWAGAGCDTSMETACGDSKNDGDGLVDCVDPCCIQLCHINPLCLGSPNPDITQE 881
 Db 782 QTWGPGGCVNAMESCADNKNEDGVLVDLPDCCLOSAQONLLCGSRDPLDIQQ 841
 QY 882 TQPVVSQQNLHFDYRIKFLVGRDSTHIIIPGENPFDGGHACVIRQVMTSDGTPLGVNI 941
 Db 842 GQ--TDWPAVKSFYDRIKLLAGKSTHIIIPGENPENSLSVLRQVVTDTGTPLVGVNV 899
 QY 942 SFVNPFLPGYTISRQDGSFDLVNCGIILIFPERAPITQHTLWLPWDRFVMTIIM 1001
 Db 900 SFVKPKGYTITRQDGTDLIANGASLTLLHFERAPFMSQBTVNLPMNSFYAMDTLVM 959
 QY 1002 RHEENEIPSCDLSNFARPNVVPSPLTSPASSCAEKGPVIVPEIOALOEISISCKMRL 1061
 Db 960 KTEENSIIPSCDLSGFRPDPIIISPLSTFFSAFQNPVPEITQVLEHEELPGSNVKL 1019
 QY 1062 SYLSRTPGYKVLRIISLTHPTIPNLMKVLWAVEGRLPKFKFAAAPDLISYFIWKT 1121
 Db 1020 RYLSRTAGYKSLKITMTQSTVPLNLRVHLWAVEGRLFKQSFQSPNLTASTFIWKT 1079
 QY 1122 DVYNOKVGLSFAFVGVYEVESCPDLILWEKRTVLQYETDASKLGWSLKHHLNI 1181
 Db 1080 DAYGRVYGLSDAVSVGFETVETCSILWEKRTALLOGLFELDPNLSGWSLKHHLNV 1139
 QY 1182 QSGILHKGNGENQFVSQQPPVITGSIINGRRRSISCPSCNGIADGNKLLAPVALTCGSDG 1241
 Db 1140 KSGILHKGNGENQFVLTQQAIIITSIMGRRRSISCPSCNGIADGNKLLAPVALAVGIDG 1199
 QY 1242 SLVYGDVNIIRIPFSGNVNITLNRKDFHSHSPAKHYLATDPMGSAVPLSDNSRR 1301
 Db 1200 SLVYGDVNIIRIPFSGNVNITLNRKDFHSHSPAKHYLATDPMGSAVPLSDNSRR 1259
 QY 1302 VFKISTVNVKDLVKNSEVWAGTQCLPFDTRCGDGKATEALTIPRGITVDFGLI 1361
 Db 1260 IYRVKSLGSKDLAGNSEVWAGTQCLPFDTEARCGDGKATDALTMSPRGFAVKNGLM 1319
 QY 1362 YFVDTGTRIRIPQNGIISTLGSNDLTSAARPLSCSDVMDSQVRLWETDALTAINPMDNSL 1421
 Db 1320 YFVDTATMRKVPQNGIISTLGSNDLTAVRELSCSDSMDVAQVRLWETDALTAINPMDNSL 1379
 QY 1422 YVLDNNVLIQISENHQVRIVAGRPVHCQVPGIDHFLSKVAIHALESATALAVSHNGVL 1481
 Db 1380 YVLENNVILRITENHQSIIAGRPVHCQVPGID--YLSKLAHSALESASAIASHTGVL 1438
 QY 1482 YIAETDEKKINRIQVTTSGEISLVAGAPSGCDKNANDCFCFSGDGVYAKDAKNTPESS 1541

Db 1439 YITETDEKKINLRQVITNGEIEICLLAGASDCDCCKNDVNCNCYSGDDAVATDAILNSPSS 1498
Qy 1542 LAVCADGELYVADLGNIRIRFIRKNEPLANTOMNYELSSPIDOELYLFDITGHHLYTOSL 1601
Db 1499 LAVAPDGTIYADJGNIRIRAVSKNKPVLNAFQYEAASPGQELYYFNADGHHQYTVSL 1558
Qy 1602 PTGDLNFTYTGDDITLITDNNGNMNVRRDSTGMPLMLVDPDGOVYVMTGTSALK 1661
Db 1559 VTGEYLYNFTYSTDNDVTELIDNNGSLKIRDSGSGMPEHLLMPDQIITLVTGTNGGLK 1618
Qy 1662 SVTTQGHLELAMWTHGNSGLLAKTSNENGWTHFYVDSFGRLLNVTFTPTQVNSFRSDTD 1721
Db 1619 VVSTQNLGLMTYDGNLTGATKSDTGTFTFYDHDHGLRNTVTRPTGVVTSLEHME 1678
Qy 1722 SSVHVQVETSSK-DDVTTITNLSAGAFYTLLOQVNSYIIGADGSLRLLLANGMEVAL 1780
Db 1679 KSIITDIENNRDDVTIVINLSSVEASYIVQDVNSYQLCNGTFLRVWYANGGISF 1738
Qy 1781 QTEPHLLAGTVNPTVGRNVTLPIDNGLNVWRQKEQARGOVTVFGRRLRVHNRNLLS 1840
Db 1739 HSEPHVLAGTITPTIGRNLISPLMENGSLNGLSEWLRKEQIKGVITFGRKLRVHGNLLS 1798
Qy 1841 LDPRVTRTEKIYDHRKEFTLRILLYDQAGPSLWSPSSRLNGVNVYSVPGYIAGTORG 1900
Db 1799 IDYDNRIRTEKIYDHRKEFTLRILLYDQAGPSLWSPSSRLNGVNVYSVPGYIAGTORG 1858
Qy 1901 MSERMEYDQAGRITSRIFADKWTSTYLEKSMVLLLSHQRYIFEFKNDRLSSVTMPN 1960
Db 1859 MSERTDIDKQRIVSRNFAKGVMSYSLDKSMVLLLSHQRYIFEFKNDRLSSVTMPN 1918
Qy 1961 VARQTLSTISVGYRNIYOPPEGNASVIOFDTEHLLHTFYLCTGRVRYKYGKLSKL 2020
Db 1919 VARHSMSTHSIGVIRNIYPPESNASVIFDSDGGRILKTSFLGTGRQVYFYKYGKLSKL 1978
Qy 2021 AETLYDITKVSFTYDETAGMLKTNINQEGFTCTIRYRQIGPLIDRQIFRFTBEGMVNAR 2080
Db 1979 SEIYVDSTAVTFGYDEITGVKXVNLQSGGFSCTIRYRKIGPLVDKQIYRFSBEGMVNAR 2038
Qy 2081 PDYNY-DNSRVTSMQAVNETPLIDLYRVDVSGTEQFGKFGVYIYDINOITAVM 2139
Db 2039 FDYTHDNSFRIASIKKIVISLPLVLYRDEISGKVEHFGKFGVYIYDINOITAVM 2098
Qy 2140 THTKHFDAYGRMKVQVEIIFRSLMYMTVQYDNNRWVKELKVPYANTTRYSYEYDAD 2199
Db 2099 TLSKHFTHGRIKVQVEMFRSLMYMTVQYDNNRWVKELKVPYANTTRYSYEYDAD 2158
Qy 2200 GOLQTSVINDKPLWRYSDYDNGNLHLLSPONSARLPLRYDIRLITRIGDVOYKMDG 2259
Db 2159 GOLQSVAVNDPRTWRYSDYDNGNLHLLSPONSARLPLRYDIRLITRIGDVOYKMDG 2218
Qy 2260 FLRORGDIPEYNSAGLLIKAYNRAGSWSRYRVDGLGRVSSKSSHLOPFYADLTN 2319
Db 2219 YLCORGSDIPEYNSAGLLIKAYNRAGSWSRYRVDGLGRVSSKSSHLOPFYADLTN 2278
Qy 2320 PTKVTHLYNHSSEITSLYYDLOGLFAMELSSGDEFYIACDNGITPLAVFSGTGLMIKQ 2379
Db 2279 PTRITHVYNSHSEITSLYYDLOGLFAMELSSGDEFYIACDNGITPLAVFSGTGLMIKQ 2338
Qy 2380 ILITYAGEIYMDTNPFOIIGHGLYDPLTKLVEMGRDDYDLVLAGRWTSPPHMLWKHL 2439
Db 2339 LQYTAAGEIYSDNPDMQVIGHGLYDPLTKLVEMGRDDYDLVLAGRWTSPPHMLWKHL 2398
Qy 2440 SSSNMVFNLYMFKNPNIPNSODIKCFMTDVSNSLLTFQGLVHNPVPGYKPDMDAMP 2499
Db 2399 GKPA-PFNLYMFKNPNIPNSODIKCFMTDVSNSLLTFQGLVHNPVPGYKPDMDAMP 2457
Qy 2500 SYELIHTOMKTQEMDNKSKSILGVOCEVQKOLKAFVTLERPDQYLGTSITSCQOAPTKK- 2558
Db 2458 PYELSESQAS-----ENQLTGTVQOQTTERRNQAFMALE-----GVITKKLHASIREKA 2507
Qy 2559 ---PASSGVSFGKGVKAFKXGRTVTDIIISVANEDGRRVAAIILNHAHLYENLHFTIDGVD 2615
Db 2508 GHWFATTTPIGKIMFAIHEGRTVTCVSSIASDSKVASVNLNAYYIDKXHSIEGKD 2567

Qy 2616 THYFVKPGPSBGLAILGLSGRRRTLENGVNVYSQINTVLNGRTRVYTDILOYGALCL 2675
Db 2568 THYFVKIGSADGLVLTGTTIGRKVLSGVNVYSQPTLLVNGRTRFTNIEFOYSTLL 2627
Qy 2676 NTRYG---TTLDEEKARVLELARQVRAQWARQOORLRGEEGLRAWTEGEKQOVLSTG 2732
Db 2628 STRYGLTPTDITLDEEKARVLDQARQALCTAWAKEQKARDREGSRLWTEGEKQOVLSTG 2687
Qy 2733 RVQGVGDGFVLSVQYPELSDSANNIHFMRQSEGR 2769
Db 2688 RVQGVGYGVYLVFVQYPELADSSNIQFLRQNMGR 2724

RESULT 12

ABG61913
ID ABG61913 standard; protein; 2725 AA.
XX
AC ABG61913;
XX
DT 15-AUG-2002 (first entry)
XX
DE Prostate cancer-associated protein #114.
XX
KW Prostate cancer; prostate tumour tissue; human; mammal; cytostatic.
XX
OS Mammalia.
XX
PN WO200230268-A2.
XX
PD 18-APR-2002.
XX
PF 12-OCT-2001; 2001WO-US032045.
XX
PR 13-OCT-2000; 2000US-00687576.
PR 08-DEC-2000; 2000US-00733288.
PR 08-DEC-2000; 2000US-00733742.
PR 24-JAN-2001; 2001US-0263957P.
PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX
WPI; 2002-471335/50.
XX
N-PSDB; ABK92230.
XX
Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
Claim 27; Page 397; 436pp; English.

XX The present invention relates to methods of detecting a prostate cancer-
XX associated transcript in a cell from a patient. The method comprises
XX contacting a biological sample from the patient with prostate cancer-
XX associated polynucleotides (designated PC genes) that selectively
XX hybridize to a sequence that is at least 80% identical to them. The
XX prostate cancer-associated polynucleotide sequences are differentially
XX expressed in prostate tumour tissue or in prostate cancer and are derived
XX from the tissues of various organisms such as humans or other mammals
XX (e.g. mice, sheep and dogs). The methods of the invention are useful for
XX diagnosing and treating prostate cancer in mammals. The prostate cancer-
XX associated genes are useful for diagnosing or treating prostate cancer,
XX as well as for identifying modulators of prostate cancer or agents that
XX inhibit prostate cancer. The nucleic acid sequences are particularly
XX useful in gene therapy, as a vaccine or in antisense applications.

```
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
SQ Sequence 2725 AA;

Query Match      63.8%; Score 9501; DB 5; Length 2725;
Best Local Similarity 62.0%; Pred. No. 0;
Matches 1737; Conservative 404; Mismatches 550; Indels 112; Gaps 17;

QY 1 MDVKERKPYRLTH-RRDAERRYTSSADSEEGKAPOKSYSSSETLTKAYDODARLAYGS- 58
DB 1 MEQTDCHPYOPLPKVKHEMLDAYTSSDESEDGRKPAQSYNSRETHYENQELRYMNSQ 60
QY 59 --RVKDIVPOEAE-EFCQRTGANFTLRELGLVEVTPPHGTLRYTDI-GLPOCGYSGMAGSD 114
DB 61 SRKRKEVEKSTQEMFECTSHILC-----SGYQTDHMSVSRHGYLEMGSD 106
QY 115 ADVEADIVLSPHPVRLWNGSTRSGRSSCLSSRANSNLITLTDTEHENTETHDGGQLQ--- 171
DB 107 VDETEGAASPDHALRMWIRGMKSEHSSCLSSRANSALSITDTHDR-KSDGNGFKFSP 165
QY 172 -----NHARLR--TTPPLLSHAHTNQHAASINSLSNENGTTPRS 209
DB 166 VCDMEQAAGSTQDVQSSPHNQFTFRPLPPPPPPHACTCARKPPPAADSLQRSMTTRS 225
QY 210 NPSAPTDHLSGEPAGGQBPAPHAQENWLLNSNIPLETRNLGKQPFLLGTLODNLIEMD 269
DB 226 QPSPA-----APAPPTSTQDSVHLHNSWVLSNIPLETR----- 259
QY 270 ILGASRDGAYSQGHFLFKQG-GTSPFCFTTSPGYPLTSTVXSPPPPLPRPTFFARPAF 328
DB 260 -----HSLFKHGSGSAIFSAASQNYPLTSTVXSPPPPLPRPTFFARPAF 305
QY 329 NLKPSKYCNWKAALSAIVISATVLLAYFVAMHLFGLNMLHQPMEGQMYEITEDTAS 388
DB 306 TFNKPYPCCNWKCTALSATAITVTALLAYVIAVHLFGLTQWLQPVVEGELYANGYSKGN 365
QY 389 SWPYPTDVSLYPSGGTGLTPDRKGGTTECKSSFPEDSFIDSGLDVGRRASQKIPP 448
DB 366 RGTESMDTTPSGKVSXSEKK-----VFQKRAIDTGEVDIGAQVMQTIPP 414
QY 449 GTFWRSQVFIIDHPVHLKFNYSGLKAALVGIYGRKGLPPSHSTQDFVELLDGRLLTQEAR 508
DB 415 GLFWRFQITTHPIYLFKNISLAKDSLGIYGRNIPTHTQDFVFKVMDGKLQVQDSK 474
QY 509 SLEGTPTQSRGTVPSPSHETGFTQYLDSGIWHLAFYNDGKESEVVSPLTTAISVNCPS 568
DB 475 GSDTQHSRNLILTSQETGFBEYMDQGPWYLAFYNDGKKYQVFLTTAISIMDDCST 534
QY 569 NCYNGDCISQTCFCLFGLQPCDGRASCPLVCSNGQYMKGRCLCHSGWKGAECQVPTN 628
DB 535 NCNGNGECISGHCHCFPGFLPCDGRASCPLVCSNGEYKXGHCVCBHGKMGCECDVPEE 594
QY 629 QCIDVACSNHGTCTGTICINPNYKGBSCBEVDCMDPTCSGRGVGVVGECHCFVWGVTN 688
DB 595 QCIDPTCFGHGTICIMGVCIQVPGYKGBICBEEDCLDPMCSNHGICVKGECCHCSTGMWGVN 654
QY 689 CETPRATLQCSGHGFLPDTGLSCDPSHTGHDSCSIEICAADCGHGVGVCGGTCECED 748
DB 655 CETPLPVQCBQCSGHGFLLDAGVCSCDPWMTGSDCSTELCTMECGSHGVCSRGICOCBE 714
QY 749 GWMGAACDQACHPRCAHEGTCRDGKCEKCPGWNHECTTIAHYLDRLVYVKEGCPGLCNGNG 808
DB 715 GWVGPTCEERSCHSHCTEHGCKDKGCEKCPGWEHGHCTTIAHYLD-AVRDGCPCGLCFGNG 773
QY 809 RCTLDLNGWHVCQGLWHRGACDTSMETACGDSKXNDGDGLVDCMDPDCCLOPLCHINPL 868
DB 774 RCTLDQNGWHVCQGVSGTGCNVVMEMLCGDNLDNDGDGLTDCVDPCCQCSQNCYSPL 833
QY 869 CLGSPNPLDIIQETQVPSQONLHSHFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGOV 928
DB 834 CQGSFDPDLIIQOQCTLFQSHTSRLFYDRIKFLIGDSTHVIPEVSFDSRACVIRGOV 893
QY 929 MTSQGTPLVGNVIGFVNNPLFGYTISRQDGSFDLVTVNGGISIILRFRAPFITQEHTLWL 988
```

```
894 VAIDGCTPLVGNVSPFLHSDYGTISRQDGSFDLVAIGGISVILIFDRSPFLPEKRTLWL 953
989 PWDREFVWETIMRHEENEIPSCDLNFAFPNPVPSPLTSPASSCAEKGPVPEIQAL 1048
954 PWNQFIVVVEKVTMQRVSDPPSCDISNFIENPIVLPSPLTSPFGSCSPERGTVIVPELQV 1013
1049 QBEISISGCKMRLSYSSLSRTPGYKSVLRISLTHPTIPENLMKVHLMAVEGRLEPRKFAA 1108
1014 QBEIIPSSFVRLSYSSLSRTFGYKTLRLALHTHSIPVGMIKVHLTVAVEGRLETKWPPA 1073
1109 APDLIYYIWKTDVYNOKVFGLSAFVSVGYEYESCPDLILWEKRTTVLQGYEIDASKL 1168
1074 AINLYVTFAWNKTDITYGQVMGLAALVSVGYEYETCPDFILWEQRTTVLQGFEDASNL 1133
1169 GWSLXDKHALNIQSGILHKGNGENQFVSQPPVGTSGIMGNRRRSISCPSCNGLADGNK 1228
1134 GWSLXKXKHLNPQSGIIHKGNGENMF:SQPPVLTIMGNHQRBSVACTWNGFAHNK 1193
1229 LLAVALVTCGSDGLYVGDVFNIRIPPSGNVTNILELRNKDFRHSHPAHKYIYLATPM 1288
1194 LFAPVALASGPDGSVVGVDFNFRIRFPISGNSVSIEL-----STSPAHKYIYLA 1246
1289 SCAPVLSDSNERRVFKISTVVVVKLVNSEVWAGTGDCLPDDTDCDGGCGKATEALT 1348
1247 SESLYLSDTNTKVKYKLSLVETKDLKSNFEVWAGTGDCLPDDTDCDGGCGKASEASLN 1306
1349 NPGRTIVDKFLIYVVDGTMIRRIDQNGIISTLGSNDLTSARPLSCDSVMDISORLEW 1408
1307 SPRGITVDRHGFIYVVDGTMIRKIDENAVITTVIGSNGLTSTQPLSCDSGMDITQVRL 1366
1409 PTDLAINPMDNSLYVLDNNVLOISENHQVRIVAGRPKHCOVPGIDHFLLSKVVAHATLE 1468
1367 PTDLAINPMDNSLYVLDNNVLOISENRRVRIIAGRP:HCQVPGIDHFLLSKVVAHATLE 1426
1469 SATALAVSHNGVLYTAETDEKKINRIROVTTSGEISLVAGAPSGCDCKNDANCDGSGDD 1528
1427 SARA-SVSHSGLLFIAETDERKVNRIQVTTNGEIIYIAGFTDCDCKIDPNCDFCSGDG 1486
1529 GYAKDAKNTSPSLVACADGELYVADLGNIRIFIRKNKPFINTQNMVELSSPIDOELYL 1588
1487 GYAKDAKMPASPSPVSPDGLYVADLGNVIRIRTSRQAHLNDNMNIYEIASFADQELYL 1546
1589 FDTTCKGHLVTSQPLTGDVLYNFTYTGDDGTLITDNNGNMNVNRDSGMPMLWLVVPGQ 1648
1547 FTWNGTHLHTLMLITRDVYVNFYNSGDLGATSSNGNSVHIRDAGMPLWLVVPGQ 1606
1649 VYVWVTMTNSALKSVTTQCHELAMMTYHNSGLLATKSNENGWTTTFYSDSFGRLTNVTF 1708
1607 VYWLITSSNGVLKRVSAQGYNPALMTYPGNTCLLATKSNENGWTTVYEDPDEGLTNATF 1666
1709 PTGQVSSFRSDTSSVHVQVETSSKDDVTITNLSASGAFYTLLODQVNSYIIGADGSL 1768
1667 PTGEVSSPHSDLEKLTKEVLDTSNRENVLMSTNLTATSTIYILKQENTQSTYRVNPDGSL 1726
1769 RILLANGMEVALQTEPHLLAGTVPNTVGNKRVNTPIDNGLNLVWQRKEQARGOVTVFG 1828
1727 RVTFASGMEIGLSSPHILAGAVNETLGKCNLSLGEHNANLIEWRKEQKNVSAFE 1786
1829 RLRLVHNRNLLSLDFDRVTRTEKIIDHRKFTLRILYDQAGRPSLWSSSRNLGNVNTYS 1888
1787 RLRLAHNRNLLSIDEDHITRTGKIYDHRKFTLRILYDQGRPILWSPVSRVNEVNTYS 1846
1889 PGGYIAGTORGITSMERMEYDQAGRITSRIFADGKTWSYTLKSVMLLHLSORVIFBPD 1948
1847 PSGLVTFIORGTWNBEMEYDQSGKIIISRTWADGKI:WSTYLEKSVMLLHLSORRIFEVD 1906
1949 KNDRLSSVTMPNVARQTLTETIRSVGYRNIYQPPENASVIOQDFTEDGHLHTFTYLGTR 2008
1907 QSDCLLSTVMPNVARHSLQTMVLSVGYRNIYTPDPSSTSFIDYSRDRGLLQTLHGTGR 1966
2009 RVLYKVKLSKLAETLYDTTKVSFYDETAGMLKINLQNEBQCTTIRYQIQLDROI 2068
```


Db 472 QYLDVGLWHLAFYNDGCKDEWVSFNTVLDSDVQDCPRNCHNGECVSGVCHCFPGFLGAD 531
 QY 592 CGRASCPVLCSGNGQVNGKCLCHSGWKAECDDVPNQCIDVACSNHGTCTGTCLNPG 651
 Db 532 CARAACPVLCSGNGQYSKGTCCQYSGWKAECDDVPNMQCIDPCSGHSGCIDGNCVCVSAG 591
 QY 652 YKESCEBVDMDPTCSGRGVCGVGRCHCFVNGGTCNCTPRATCLDQCDSHGHTPLPDG 711
 Db 592 YKGEHEBVDCLDPTCSSHGVCVNGECLCSPGWGLNCELARVQCPDQCSHGHTVLPDGT 651
 QY 712 LCSDFSWTGHDCSIBICAAOCGGHVCVCGTCTCEDGWMGAACDQACPRCAEHGTCR 771
 Db 652 LCSDFENWMPDCSV----- 666
 QY 772 DGRCECSPGWNGHCIIAHYLDVWVKEGCPGLCNGNGRCTLDLINGVHCVCOLGWRGAGCD 831
 Db 667 -----DGCPLCNGNGRCTLGQNSQVCCQCTGWRGPGCN 700
 QY 832 TSMETAGBDSKONDGLVDMDPCDCLQPLCHINPLCLGSPNPLDIIQETVPVSSQNL 891
 Db 701 VAMETSADNKNEDGLVDCLDPCDCLQACQNSLLCRGRDPLDIIQCGQ--TDWPAV 758
 QY 892 HSFYDRIKFLVGRDSTHIIIGENPFDGCHACVIRGOVMTSDGTPLVGNMISFVNNPLFCY 951
 Db 759 KSPYDRIKLAGDSTHIIIGENPFNSLSVLRGOVMTDGTPLVGNVNSFVKYPRGY 818
 QY 952 TISRQDGSPLVNGGISILIRPERAPFITQEHTLWLPMDRFFVMTETIWRHEENEIPSC 1011
 Db 819 TIRQDGTPLIANGASLTLHERAPFMSQERTVLPWNSFYAVDTLVAKTEENSIPSC 878
 QY 1012 DLNFRAPNPVSPSPITSFASCAEKGPIVPIQALQBEISISGCMRLSYLSSTPGY 1071
 Db 879 DLGQFVRPDIITSSPLSTFSAAPQNPITVPETQVLHBEIEPLPGSNVKVLRYSRTAGY 938
 QY 1072 KSVLRSLTHPTIPENLMKVLHVAVEGRPLFRKFAAAPLSYFTWKTVDVNOXVFL 1131
 Db 939 KSLKLTWQSTVPLNLRVHLMVAVEGHLFQKSFQASPNLAYFTWKTDAYQORVYGL 998
 QY 1132 SEAFVSYGYBESCPDILWEKRTVLQGYEIDASKLWGLSKDHALNIQSGILHKGNG 1191
 Db 999 SDAVSVGYEYETCPSLILWEKRTALQGFELPDSNLGWSLKDHALNVKSGILHKGNG 1058
 QY 1192 ENQFVSQOPVIGSIMGRRRSITSCPSNGLADGNKLIAPVALTCGSDGLVVGDFNYI 1251
 Db 1059 ENQFLTQOPAITITSIMGRRRSITSCPSNGLAEGNKLIAPVALAVIDGSLVVGDFNYI 1118
 QY 1252 RRIFFPSGNVTNILELRNKRFRHSHSPAKHYLATDPMGSAVFLSDNSRRVFKIKSTVVY 1311
 Db 1119 RRIFFPSRNVTSILELRNKEFKHSNNPAHYLLAVDPVSGSLYVSDTNSRRIYRVKSLGT 1178
 QY 1312 KOLVNSEVAVAGTGCQCLPDDTRCGDGGKATEATLTNPRGITVDFKGLIYFVDTWIR 1371
 Db 1179 KDLAGNSEVAVAGTGCQCLPDEARCGDGGKAIDATLMSPRGIAVDKNGLMYFVDATWIR 1238
 QY 1372 IDONGIITSLIGSNDLTASRPLSCDSYNDISQVLEMPDPLAINPMDSNLYVLDDNNVLQ 1431
 Db 1239 VDQNGIITSLIGSNDLTAVRPLSCDSNDVAQVLEMPDPLAINPMDSNLYVLDDNNVL 1298
 QY 1432 ISENHQRIVAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKI 1491
 Db 1299 ITENHQVSIILAGRPMHCQVPGID--YSLSKLAITHSALESASAIATSHGTGLYIETDEKKI 1357
 QY 1492 NRIRQVTSIGSILVAGAPSCDCNDANDCFSGDDGYAKADAKLNPSSLSIACAGELY 1551
 Db 1358 NLRQVTTNGSICLLAGAASDCDKNVNCYSGDDAYATDAILNPSLSIACAGELY 1417
 QY 1552 VADLGNIRIRIRKKNKPLNTQNNWELSSPIDQELYLFDPTGKHLYTQSFTGDLYNFT 1611
 Db 1418 IADLGNIRIRAVSKNPVLAFAQYEAASPEQEELYVFNADGIHQYVTSVLTGEYLYNFT 1477
 QY 1612 YTGDCDITLITDNNGNMNVRRDSTGMPLMLVDPDQGVYVMTGMTNSALKSVITQGHCLA 1671
 Db 1478 YSTDNDVTELDNNGNSILKIRRDSGMPRHLIMPDNQIITLTVTGNGGLKVWSTQNLELG 1537

QY 1672 MMTYHNSGLLATKSNENGWTTTFEYDSFGSLTNVTPTGOVSSFRSDTSSVHVQVETS 1731
 Db 1538 LMTYDGNITGLLATKSDTGTWTTFYDDBEGELTNVTPTGVVTSLSHREMEKSIITIDIENS 1597
 QY 1732 SK-DDVITITNLSAGAFYTLLOQVNRNSYIIGADGSLRLLANMEVALOTEPHLAGT 1790
 Db 1598 NRDDVITITNLSVSEASYTVQOVNRNSYQLCNNGTLRVMYANGMISFSEPHVLAGT 1657
 QY 1791 VNPVTGKENVTLPTDNGLNLVEMRQEXQAGOVTVFGRRLRVHNRNLLSLDFORVTRTE 1850
 Db 1658 IPTTIGRNCISLPMENGLNLSIEWLRKEQIKGKVTIIFGRKLVRGRNLLSIDYDNRITE 1717
 QY 1851 KIYDHRKFTLRIILYDQAGRPSLSPSRNLGNVNTYSPGGYIAGIORGIMSEMEVDQA 1910
 Db 1718 KIYDHRKFTLRIIYDQVGRPLMLPSSGLAAVNVSYFFNGRLAGLQAGMAMERTIDKQ 1777
 QY 1911 GRITSRIPADOKTWSYTYLEKSMVLLLSHQRYIFEDKNDRLSSVMNPNVAROTLEIR 1970
 Db 1778 GRIVSRMFADOKWWSYLDKSMVLLLSHQRYIFEDSSDRLLAVTNPSVARHSMSTHT 1837
 QY 1971 SVGYRNTYQPEGNASVIOQFTEDGHLHLLFTYLGTRRVYKYGKLSKLAETLYDTTKV 2030
 Db 1838 SIGYRNTYNPESNASVIFDYSDGRLLKTSFLGTGRQVYKYGKLSKLSSEIVYDSTAV 1897
 QY 2031 SFTYDGTAGMLKTNLQNEGTCTIIRYQIQLDRQIFRTEGMMVNAEDYNY--DNSF 2089
 Db 1898 TFGYDETTGLVMKNVLSQSGFSCTIRYKIGPLVDKQIYREFSEEGMMVNAEDYTHDNSF 1957
 QY 2090 RYVSMQAVINETPLDLYRYDDVSGKTEQFGKGVYIYDINOIITTAVMTHTKHFDAYG 2149
 Db 1958 RIASKPVISETPLVDLYRYDESGKVEHFKGVYIYDINOIITTAVMTHSKHFDHNG 2017
 QY 2150 RMKEYQYBIFRSLMYMTVQYDNMGRVYKELKGVYANTTRYSEYDADQOLQTVSIND 2209
 Db 2018 RIKEYQYEMFRSLMYMTVQYDSMRVYKRELKGPYANTTRYKYDYDGDQOLQSVAYND 2077
 QY 2210 RPLWYSYDINGNLHLLSPGNSARLTPLRYDIRITELGVOYKMDDEGFLRORGDIIF 2269
 Db 2078 RPTWYSYDINGNLHLLSPGNSVRMLPLRYDLRITRLGDOYKIDDDGFLCQSGDIIF 2137
 QY 2270 EYNSAGLLIKAYNRAGSVRYRYDGLGRRVSSKSHSHLQFFYADLTNPKTVTHLYNH 2329
 Db 2138 EYNSKGLLTRAYNKASGWSQYRYDGVGRRASYKTNLGHLLQYFYSDLHNPTRITHVYNH 2197
 QY 2330 SSESITSLYDLOGLHLPAMELSSGDEFYACDNTGTPLAVPSGCTGLMKIQLITYAGELY 2389
 Db 2198 SNEITSYDLOGLHLPAMESSGSEYVYASDNTGTPLAVFSINGLMTIKQOYTAIGEYI 2257
 QY 2390 MDTNENFOLIIYGHGGLYDPLTKLVHMGRRDDYDLVLAGRWTSPDHKLWKLSSNNVMPNL 2449
 Db 2258 YDSNDFQVQVIGFHGGLYDPLTKLVHFTQDYDVLVAGRWTSPDYTMKNVKGKEPA--PNL 2316
 QY 2450 YMFKNPNISQDICKFMTDVNSWLLTFGFOLHNVIPGPKPDMADMEPSVELIHTOMK 2509
 Db 2317 YMFKNPNLSSELDKYNVTDVKSMLVMFGFOLSIIIPGFPRAKMYFVPPPELSESOAS 2376
 QY 2510 TQEWDSKSLILGQVCEQVKOLKAFVTLERFDOLYSTITSQCAQPKTKK---PASSGSV 2565
 Db 2377 ----ENGOLITVQVQTERHNOAFWALE-----QVITKKUHASIREKAGHFWATTPI 2426
 QY 2566 FGKGVKFPALKGRVTTDIIISVANEDGRRVAAILNHAHYLENLHFTIDGVYDTHFYVKPGPS 2625
 Db 2427 ICKGIMFALKEGRTVITGVSSIASDSRKVASVANNAYVLDKMHYSIEGDKTHFYVKIGSA 2486
 QY 2626 EGDAILGUSGGRRTLENGVNTVTSQINTVLNGRTRRTIDIOLOVGCALCLNTRYG---TT 2682
 Db 2487 DGDVLTGTTIGKRVLESQVNTVTSQPTLLVNGRTRRTNIEFQYSTLLSIRYGLTPTD 2546
 QY 2683 LDEKARVLELQARAVROAWAREOORLREGEGERAWTEGKQOVLSTGRVQVQDGFVF 2742
 Db 2547 LDEKARVLDQARQALGTAWAKEQOKARDGREGSELWTEGKQQLSTGRVQVGEYV 2606

Db	1907	QSDCLLSVTWPSMVRHSIQTMLSVGYRNITPDSSTSFIOQYSRDGRLLQTLHLGTGR	1966
Qy	2009	RVIYKYGKLSKLAETLYDTTKVSTFYDETAGMLKTNLQNEGFTCTIRYQIGPLIDROI	2068
Db	1967	RVLKYTKQARLSEVLYDTTQVTLTYBESSGVIXIHLMDGFTCTIRYQIGPLIDROI	2026
Qy	2069	FRFTEEGMVARFDYNDNSPRVTSMOAVINETPLPIDLYRDDVSGKTEQFGKFGVIYY	2128
Db	2027	FRFSEEGLVNARFDYSY-NNERVISMQAVINETPLPIDLYRYVDVSGRTEQFGKFSVINY	2085
Qy	2129	DINQIITTAVMTHTKHFDAYGRMKVEQYEIFRSLMYMWTVOYDNMGR	2175
Db	2086	DLNQVITTTVMKHTKIFSANGQVIEQVEILKALAYWMTIQYDNVGR	2132

Search completed: June 24, 2004, 16:19:25
Job time : 105 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 16:15:58 ; Search time 33 Seconds
(without alignments)
8071.343 Million cell updates/sec

Title: US-10-029-020-14
Perfect score: 14887
Sequence: 1 MDVKERKPYRSLTRRDAER.....ELSDSANNHFRQSEMGRR 2769

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14269	95.8	2825	2 T14271	Doc4 protein, stre
2	3804.5	25.6	2515	2 S47008	tenascin-like prot
3	3515.5	23.6	2406	2 A54148	odz protein - frui
4	3111.5	20.9	849	2 T46253	hypothetical prote
5	2109	14.2	2531	2 T16743	hypothetical prote
6	1175.5	7.9	782	2 A61625	tenascin-like prot
7	943	6.3	184	2 T12457	hypothetical prote
8	659.5	4.4	2019	1 JQ1322	tenascin precursor
9	652	4.4	1746	1 S19694	tenascin precursor
10	649.5	4.4	2201	2 A32160	tenascin-C - human
11	641	4.3	1810	1 A32230	tenascin precursor
12	626	4.2	4006	2 T09070	probable tenascin
13	592	4.0	3566	1 A40701	tenascin-X precurs
14	577.5	3.9	647	2 A43902	tenascin - eastern
15	572.5	3.8	4135	2 T42629	tenascin-X - bovin
16	438.5	2.9	2703	1 A24420	notch protein - fr
17	429	2.9	2524	2 A35844	notch protein - Af
18	428.5	2.9	2555	2 A40843	notch protein homo
19	427.5	2.9	2531	2 A46019	notch-1 protein -
20	424	2.8	1203	2 A49175	Notch B protein -
21	423.5	2.8	1064	2 A40136	fibropellin Ia - s
22	422.5	2.8	2531	2 S18188	notch protein homo
23	422	2.8	1620	2 T27283	hypothetical prote
24	420	2.8	2471	2 A49128	cell-fate determin
25	419.5	2.8	1220	2 A56136	jagged protein pre
26	413	2.8	2437	2 T22612	transmembrane prot
27	410	2.8	3191	2 T22945	hypothetical prote
28	400.5	2.7	1295	2 A32901	glp1 protein precu
29	399	2.7	2531	2 T31070	notch homolog - se

30	392.5	2.6	1353	1 JH0675	restrictin precurs
31	392.5	2.6	2321	2 S78549	notch3 protein - h
32	389	2.6	1574	2 T13954	MEGF6 protein - ra
33	387.5	2.6	1722	2 B89753	protein FilC7.4 [1
34	387.5	2.6	1964	2 T09059	notch4 - mouse
35	385.5	2.6	1356	2 A45445	janusin precursor,
36	384	2.6	2318	2 S45306	notch 3 protein -
37	383	2.6	1408	2 S16148	gene serrate prote
38	382.5	2.6	2139	2 S15672	crumbs protein - f
39	382	2.6	2352	2 T30201	Notch homolog prot
40	381	2.6	861	2 A48825	Notch homolog Notc
41	373.5	2.5	833	2 S19087	gene Delta protein
42	372.5	2.5	832	2 A31246	neurogenic protein
43	372.5	2.5	880	2 S00670	neurogenic repetit
44	369.5	2.5	1111	2 T26972	hypothetical prote
45	356.5	2.4	473	2 A56175	adhesive plaque pr

ALIGNMENTS

RESULT 1

Ti4271
Doc4 protein, stress-induced - mouse
N:Alternate names: odz protein homolog
C:Species: Mus musculus (house mouse)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14271
R:Wang, X.Z.; Kuroda, M.; Sok, J.; Batchvarova, N.; Kimmel, R.; Chung, P.; Zinszner, H.
EMBL J. 17, 3619-3630, 1998
A>Title: Identification of novel stress-induced genes downstream of chop.
A:Reference number: 217951; MUID:98315054; PMID:9649432
A:Accession: T14271
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2825 <WAN>
A:Cross-references: EMBL:AF059485; NID:g3170614; PID:g3170615; PIDN:AAC31807.1
C:Genetics:
A:Gene: Doc4

Query Match	95.8%	Score 14269;	DB 2;	Length 2825;
Best Local Similarity	93.4%	Pred. No. 0;		
Matches 2654;	Conservative 47;	Mismatches 52;	Indels 88;	Gaps 4;
QY	1	MDVKERKPYRSLTRRDAERRYTSSADSEEGKAPQKSYSSSETLKAYDQARLAYGSRV	60	
Db	1	MDVKERKPYRSLTRRDAERRYTSSADSEEGKAPQKSYSSSETLKAYDQARLAYGSRV	60	
QY	61	KDIVPOEAEEFCRTGANFTLRELGLEVTTPHGTLYRTDGLPCGYSMGAGSDADMEAD	120	
Db	61	KDMVPOEAEEFCRTGNTFTLRELGLGEMTPPHGTLYRTDGLPCGYSMGAGSDADLEAD	120	
QY	121	TVLSPHEFVRLWGRSTRSGRSCLSRANSNLTLTDEHENTET	164	
Db	121	TVLSPHEFVRLWGRSTRSGRSCLSRANSNLTLTDEHENTETGAPLHCSSASSTPIEQ	180	
QY	165			DHPGGL 170
Db	181	SPSPPPPPANESORRLNGVQAFTPDSDSEEFVFNPSFLVKSGSASLGVAANDHPSL	240	
QY	171	QNHAFRLRTPPPLGSHAHTPNQHHAAASINSLNRGNFTPRSNPSAPTDHLSGEPAGGAQ	230	
Db	241	QNHFLRLRTPPPLPHANTPNQHHAAASINSLNRGNFTPRSNPSAPTDHLSGEPAGSAQ	300	
QY	231	EPAAQENWLNLSNIPLETRNLGKOPFGTLODNLMIEMDIIGASRHGAYSDGHFLFKPG	290	
Db	301	EPHAQDNWLNLSKIPVETRLGKOPFLGTWQDNLMIEMDIIFSARRRDGAYSDGHFFFKPG	360	
QY	291	GTSPLFCCTTSQGYPLTSSTVYSPPRPLPRSTFARPAFNLKKPKSKYCNWKAALSAIVIS	350	
Db	361	GTSPLFCCTTSQGYPLTSSTVYSPPRPLPRSTFARPAFNLKKPKSKYCNWKAALSAIIS	420	
QY	351	ATLVILLAYFVAMHLFGLNWLQPMEG--QMYEITEDTASSWPVPTDVSLYPSGSGTGLT	408	

421 ATLVLAVFVAMEFLGLNHLQPMGQOMYEITEDTASSVPVPTDVSILPSGGTGLT 480
 409 PDRKGGKTTGKSPSSFPEDSFDSGEIDVGRASQKIPPGTFWRQOVFIDHPVHLKFN 468
 481 PDRKGAABGKSPSSFPEDSFDSGEIDVGRASQKIPPGTFWRQOVFIDHPVHLKFN 540
 469 SLGKAALVGIYGRKGLPPSHSTQDFVELLDGRLLTQEARSLGCTPRQSRGTVPSSSHET 528
 541 SLGKAALVGIYGRKGLPPSHSTQDFVELLDGRLLTQEARSLGCTPRQSRGTVPSSSHET 600
 529 GFTQYLDGSIWHLAFYNDGKESVWFLTATBSVDNCPSCNVCNGDCTSGTCHFLGFL 588
 601 GFTQYLDGSIWHLAFYNDGKESVWFLTATBSVDNCPSCNVCNGDCTSGTCHFLGFL 660
 589 GPCGRASCPLCSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSHGTCITGTCIC 648
 661 GPCGRASCPLCSGNGQYMKGRCLCHSGWKGAECVPTNQCIDVACSHGTCITGTCIC 720
 649 NPGYKGESCEVDCMDPTCSGRGVVRGCHCFVGMGGTNCETPRATCLDQCSHGTFPL 708
 721 NPGYKGESCEVDCMDPTCSGRGVVRGCHCFVGMGGTNCETPRATCLDQCSHGTFPL 780
 709 DTGLCSDPSTWGHDCSIEICAADCGHGVVCGTCCEDGWMGAACDORACHPRCAEHG 768
 781 DTGLCSDPSTWGHDCSIEICAADCGHGVVCGTCCEDGWMGAACDORACHPRCAEHG 840
 769 TCRDGCESCPGNGEHCTIAHYLDRAVKEGCPGLCNGNGRCTLDLNGWHVCVQLGWRGA 828
 841 TCRDGCESCPGNGEHCTII-----EGCPGLCNGNGRCTLDLNGWHVCVQLGWRGT 891
 829 GCDTSMETACGSKNDGDLVDCMDPDCCLQPLCHINPLCLGSPNPLDIIQETQVPVSQ 888
 892 GCDTSMETGCGDKNDGDLVDCMDPDCCLQPLCHVNPCLGSPDPLDIIQETQAPVSQ 951
 889 QNLHNSFYDRIKFLVGRDSTHIIIPGENPFDGHCACVIRGQVMTSDGTPPLVGVNIFSWNPL 948
 952 QNLHNSFYDRIKFLVGRDSTHIIIPGENPFDGHCACVIRGQVMTSDGTPPLVGVNIFSWNPL 1011
 949 FGTYISROGSPDLVNGGSIILPERAPFITQEHHTLWLPDRFFVMTIIMRHEENI 1008
 1012 FGTYISROGSPDLVNGGSIILPERAPFITQEHHTLWLPDRFFVMTIIMRHEENI 1071
 1009 PSCDLNENARPVPWSPSLTPASSCAEKGPVPEIQALQEEISISGCKMELSVLSRT 1068
 1072 PSDDLNFARPNVPWSPSLTPASSCAEKGPVPEIQALQEEISISGCKMELSVLSRT 1131
 1069 PGYKSVLRISLTHPTIPFNLKVMVAVVEGRFLFRKWFAPADLSYFIFWDXTDYINQKV 1128
 1132 PGYKSVLRISLTHPTIPFNLKVMVAVVEGRFLFRKWFAPADLSYFIFWDXTDYINQKV 1191
 1129 FGLSEAFVSVGYEYSCPLIILWEKTTVLQGYEDASKLGWSIDKHALNIQSGILHK 1188
 1192 FGLSEAFVSVGYEYSCPLIILWEKTTVLQGYEDASKLGWSIDKHALNIQSGILHK 1251
 1189 GNGENOFVQPPVIGSIMGNGRRRSISCPSCNGLADGNKILAPALTCGSDGSLYVGDF 1248
 1252 GNGENOFVQPPVIGSIMGNGRRRSISCPSCNGLADGNKILAPALTCGSDGSLYVGDF 1311
 1249 NYIRIRFPNGVNTNILELRNKRFRSHSPAHKYIATDPMGAVFLSDNSNRFPKIKST 1308
 1312 NYIRIRFPNGVNTNILEM-----SHSPAHKYIATDPMGAVFLSDNSNRFPKIKST 1364
 1309 VVVKDLVKSSEVAVAGTQDCLPFDDTRCGDGKATEATLTNPRGTVDKFGLIYFVDCGM 1368
 1365 TVVVKDLVKSSEVAVAGTQDCLPFDDTRCGDGKATEATLTNPRGTVDKFGLIYFVDCGM 1424
 1369 IRRIDQNGIISTLLGSDNLTSAARPJSCDSVMDISQVLEWPTDLAINPMSNLSVYLDNNV 1428
 1425 IRRVDQNGIISTLLGSDNLTSAARPJSCDSVMDISQVLEWPTDLAINPMSNLSVYLDNNV 1484
 1429 VLQISENHQVRIVAGRPMHCQVPGIDHFLSKVAITHATLESATALAVSHNGVLYIAETDE 1488

1485 VLQISENHQVRIVAGRPMHCQVPGIDHFLSKVAITHATLESATALAVSHNGVLYIAETDE 1544
 1489 KKINRIQVITTSGBEISILVAGAPSCDCKNDANDCDSGDDGYAKADKINLPSSILVACADG 1548
 1545 KKINRIQVITTSGBEISILVAGAPSCDCKNDANDCDSGDDGYAKADKINLPSSILVACADG 1604
 1549 ELYVADLGNIRIRIRKKNKPLANTQNMVELSSPIDQELYLFTDTSGRKLYTQSLPTGDYLY 1608
 1605 ELYVADLGNIRIRIRKKNKPLANTQNMVELSSPIDQELYLFTDTSGRKLYTQSLPTGDYLY 1664
 1609 NFTVTGGDITLITDNNGNMNVRRDSTGMPMLVVPDQGVYVWTGNSALKSVTTQGH 1668
 1665 NFTVTGGDITLITDNNGNMNVRRDSTGMPMLVVPDQGVYVWTGNSALKSVTTQGH 1724
 1669 ELAMWTVHGSGLLATKSNENGMWTFYEDSPGRLTNVTPPTQOVSSFRSDTSSVHVQV 1728
 1725 ELAMWTVHGSGLLATKSNENGMWTFYEDSPGRLTNVTPPTQOVSSFRSDTSSVHVQV 1784
 1729 ETSSKDDVTITTLNLSASGAFYTLQDOVRNSYIIGADGSLRLLLANGMEVALQTEPHLLA 1788
 1785 ETSSKDDVTITTLNLSASGAFYTLQDOVRNSYIIGADGSLRLLLANGMEVALQTEPHLLA 1844
 1789 GTVNPVTGKRNVTLPIDNGNLVWQRKEQARGQVTFGRRLRVNRNLLSLDFDRVTR 1848
 1845 GTVNPVTGKRNVTLPIDNGNLVWQRKEQARGQVTFGRRLRVNRNLLSLDFDRVTR 1904
 1849 TEKIDYDHRKFTLIRILYDQAGRPFLSPSSRLANGVNVYSPGGYIAGIQRGIMSERMEYD 1908
 1905 TEKIDYDHRKFTLIRILYDQAGRPFLSPSSRLANGVNVYSPGGYIAGIQRGIMSERMEYD 1964
 1909 QAGRITSRIIFADGKTWSYTYLEKSMVLLHSQRYIFEPDKNDRLSSVTMPNVAROTLET 1968
 1965 QAGRITSRIIFADGKTWSYTYLEKSMVLLHSQRYIFEPDKNDRLSSVTMPNVAROTLET 2024
 1969 IRSVGYRNIYQPPGEGNASVIOQFTEDGHLHHTFGLTGRVRYKYGKLSKLAETLYDIT 2028
 2025 IRSVGYRNIYQPPGEGNASVIOQFTEDGHLHHTFGLTGRVRYKYGKLSKLAETLYDIT 2084
 2029 KVSETYDTHAGMLKTINLQNEGTCTIRYQIOPIDROQIFRTEBGMNARFDYNDNS 2088
 2085 KVSETYDTHAGMLKTINLQNEGTCTIRYQIOPIDROQIFRTEBGMNARFDYNDNS 2144
 2089 FRVTSQAVINETPLPDLRYDDVSKTEQFGKFGVYIYDINOIITTAWTHSKLFDAY 2148
 2145 FRVTSQAVINETPLPDLRYDDVSKTEQFGKFGVYIYDINOIITTAWTHSKLFDAY 2204
 2149 GRMKEVOYEIFRSLMWTMTVOYDNMGRVVKELKVGVPYANTTRYSEYDADGQLOTVSIN 2208
 2205 GRMKEVOYEIFRSLMWTMTVOYDNMGRVVKELKVGVPYANTTRYSEYDADGQLOTVSIN 2264
 2209 DKPLWRYSDNLGNLHLISPGNSARLTPLRYDTRDRIITLGDVQYKMDGDFLROGGDI 2268
 2265 DKPLWRYSDNLGNLHLISPGNSARLTPLRYDTRDRIITLGDVQYKMDGDFLROGGDI 2324
 2269 FEYNSAGLLIKAYNRAGSMVRYDYDGLGRVRSKSSSHSHLQFFYADLTNPVKVTHLYN 2328
 2325 FEYNSAGLLIKAYNRAGSMVRYDYDGLGRVRSKSSSHSHLQFFYADLTNPVKVTHLYN 2384
 2329 HSSSEITSLYDIOGHLPAMELSSGDEFYIACDNIGTPLAVSGTGLMTKQILYTAIGE 2388
 2385 HSSSEITSLYDIOGHLPAMELSSGDEFYIACDNIGTPLAVSGTGLMTKQILYTAIGE 2444
 2389 YMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDVLVLAGRWTSPDHMLWKHLSSSNWMPFN 2448
 2445 YMDTNPNFQIIIGYHGGLYDPLTKLVHMGRRDVLVLAGRWTSPDHMLWKHLSSSNWMPFN 2504
 2449 LYMPKNNPISNSQDIKCFMTDVSMLTFPGPOLHNVIPGYPKPMDMAMEPSYELIHTOM 2508
 2505 LYMPKNNPISNSQDIKCFMTDVSMLTFPGPOLHNVIPGYPKPMDMAMEPSYELIHTOM 2564
 2509 KTOEWNDSKILGVQCEVQKQKAFVTLERFDOLVGTSTISCCQAPKTKKFASSGSPFGK 2568
 2565 KTOEWNDSKILGVQCEVQKQKAFVTLERFDOLVGTSTISCCQAPKTKKFASSGSPFGK 2624

QY 2569 GVKFALKDGRVTTDIISVANEDGRRVAAIILNHAHYLENLHFTIDGVDTHYFVKPGPSBGD 2628
DB 2625 GVKFALKDGRVTTDIISVANEDGRRVAAIILNHAHYLENLHFTIDGVDTHYFVKPGPSBGD 2684
QY 2629 LAILGLSGGRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYCALCLNTRYCTTIDDEKA 2688
DB 2685 LAILGLSGGRTLENGVNVTVSQINTVLNGRTRRYTDIQLQYCALCLNTRYCTTIDDEKA 2744
QY 2689 RVLELARQAVQAWAREQORLEBEGELRAWTEGKQOVLSTGRVQDGGFFVLSVEQY 2748
DB 2745 RVLELARQAVQAWAREQORLEBEGELRAWTDGKQOVLSTGRVQDGGFFVLSVEQY 2804
QY 2749 PELSDSANNIHFMRQSEMGRR 2769
DB 2805 PELSDSANNIHFMRQSEMGRR 2825

RESULT 2
S47008
tenascin-like protein - fruit fly (*Drosophila melanogaster*)
C:Species: *Drosophila melanogaster*
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 21-Jul-2003
C:Accession: S47008
R:Baumgartner, S.; Martin, D.; Hagios, C.; Chiquet-Ehrismann, R.
EMBO J. 13, 3728-3740, 1994
A:Title: Ten(m), a *Drosophila* gene related to tenascin, is a new pair-rule gene.
A:Reference number: S47008; MUID:94349920; PMID:8070401
A:Accession: S47008
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-2515 <BAU>
A:Cross-references: EMBL:X73154; NID:G510505; PIDN:CAA51678.1; PID:G510506
C:Genetics:
A:Gene: FlyBase:Ten-m
A:Cross-references: FlyBase:FBgn0004449
F:298-324/Domain: EGF homology <EGF>
F:394-421/Domain: EGF homology <EGF1>

Query Match 25.6%; Score 3804.5; DB 2; Length 2515;
Best Local Similarity 33.2%; Pred. No. 4.9e-203;
Matches 818; Conservative 497; Mismatches 969; Indels 181; Gaps 54;

QY 384 EDTASSWPVTDVSLYPSGGTGLTDPDRKKGKTTEKPSFPEDSFIDSGIDVGRAS 443
DB 131 EATSSAATSSQSLSLTPSLSSSLANANNGARTF---PARSPPDG-TTFGQITLGGKLT 186
QY 444 QKTPPGTFWRQVFDHPVHLKENVSLGKAALVGIYGRKGLPSPHQDFVLLDGRLL 503
DB 187 KEIQPSYNNMQYQSEPAVKEFYTPRGASIGVYGRNALPHTQYHFKEVLSGFSAS 246
QY 504 TQEAR--LEGTPQSGRTVPPSSHETGFIQYLDGIMHLAFYN-DGKSESVVSPFTTAT 560
DB 247 TRTAAHLISITREVR-----YMEPGHWFVSLYNDGDVQELTFVAVAE 292
QY 561 ESDVNCPSNYGNGDCISGTCFLGLPGDCCGRASCPVLCSGNGOYMKGRCLCHSGWK 620
DB 293 DMTQNCPSNGGQCLLHCQCNPFGGDDCESVCPVLCSQGHETNGECINPGWK 352
QY 621 AECDFVTNQCIDVACSNHGTCITGTCINPFGKESCEEVDCMDPTCSGRGVGVRGECH 680
DB 353 KECSLRHDECEVADCSGHGCVSKQCQVRYGKFCCEEVDPCPNCSGHGFCADGTCIC 412
QY 681 FVWGGTNCT---PRATCLDQCSGHGTFLPDGLCSDPSWTGHCDSIEICNADCGHG 737
DB 413 KKGWGPDCATMDQDALQCLPDCSGHGTFLDPTQTCTCAKWSGDDCSKELCDLDCGQHG 472
QY 738 VCVGGTCRCRDGWMGAACDQACPRCAEHGTCRDGKCECSPWNGEHCCTIAHYLDRVVK 797
DB 473 RCEGDACADPEWGEYCNTRLCVRCNEHGQCKNGTCLCVGNGKCHCTI----- 523
QY 798 EGGFGLCNGNGRTLDLNG-WHCVCQIGWRGAGCDTSMETACGDSKDNKDGLVDCMDPD 856

DB 524 EGCPSAGHCQCRVSGEGQWECRCYEGWDGPDGCGIALELNCGSKDNKDGLVDCEDPE 583
QY 587 CCLQPLCHINPLCLGSPNPLDIQETQVPVYQQNLHSHFYDRIKELVSGDSPHIIPGENPF 916
DB 584 CCASHVCKTSLCVSAKPIDVLRKQPPAITA---SFFERMKELIDESSLQNTAKLET 640
QY 917 DGGHACVIRGQVMTSDGTPLVGNISFVNNPFLFGYTTISROQDPSDLVNGSGISILRFR 976
DB 641 NESRSARIRGRVTVSLGVLGVGRVS-TTTLLEGFTLTRDDGWFDMVNGGAVTLOFGR 699
QY 977 APFTTQHTLWLPDRFFVMEIIMRHEHE-----IPSCDLSNFARPNVPSPLTS 1030
DB 700 APFPQSRIVQVPNVEVIIDLVMMSSEKGLAVTTHTCFADHYDLMKPVLWASWKHG 759
QY 1031 FASSCAEKGRPIVPEIQALQEBEISISGCKMRLSYLSSTTPGYKSVLRISLTHPTIPFNLK 1090
DB 760 PQGACPDRESAILAESQVIESLOIPGTLNLVHSSRAAGVLSIKLQLPDVIPTSLHL 819
QY 1091 VHLMAVEGRLEFRKWFAPAAAPDSLYFIWDTVDVYNQKVLGSEAFVSGVEYESCPDLIL 1150
DB 820 IHLRITIEGILFERIFEADPGIKFTYANNRLNIYQRVGVTTAVWVGQYTDCTD-IV 878
QY 1151 WEKTTVLQGYEIDASKLGWSLKKHALNIQSGILHKGNGENOFVSQQPVGISIMNG 1210
DB 879 WDIQTTLSGHDSISEVGGWNLDIHRYNPHFEGILOKSGSNYLRNKPRIILUTWGDG 938
QY 1211 RRRSISPCSNGLADGNKLAPVALTCGSDGSLYVGFNFYIRIRFPSSGNVTNILELRNKD 1270
DB 939 HORPLECPDCDQATKORLLAPVALAAAPDGLSVGDFNVIIRIWTGSIINTVVKL--- 994
QY 1271 FRHSHSAHYKYLATDPMGSAVFLSDNSRVPFKISTVWVVDLVKNSVAVAGTDCQCLP 1330
DB 995 --NATRSYRYHMAISPLDGLTYVSDPESHQIIEVRDINDYVQPELWAEVVGGERCLP 1052
QY 1331 FDDTRCGDGGKATEATLTPRGITWDXFGLIYVDGTMIRRIDQNGIISTILLGNDLTS 1390
DB 1053 GDEAHCGDGLAKADAKLAYPKGIAISSDNILYFADGTINRMVDRDGLVSTLIGNEHKS 1112
QY 1391 -RPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDDNNVLQISENHQVRIVAGRPHCQ 1449
DB 1113 WKPIPCGCTKLEEMHLRWPTLAVSPMDNTHIIDHMLRMTPDGRVVRVISGRPLHCA 1172
QY 1450 VPGIDHFLSKVATHLESATALAVSHNGVLYIAETDEKKINRIROVTSSETSLVAGA 1509
DB 1173 TASTAYD--TDLATHAILVMPQSIAPGLGELYVAESDSQRINRVIGTGRAPPAGA 1230
QY 1510 PSGCDCKNDANCDCFGDDGYAKDAKLNTFSSLAVCADGELYADLGNIRIRIRKPKPF 1569
DB 1231 BSKCNCL-ERGDCDCEAEHYLATSAKENTIAALAVTPDSHVHIAADQANRIRSVMSSTPE 1289
QY 1570 LNTQNMVELSPIDQELVLPDTTCKHLYTQSLPTGDLVLYNFTY---TCDGDTILITDNG 1626
DB 1290 ASPGREYBIYAPDMQEIYIFNRFQGHVSTRNLTGETTYVFTYNNVNTSNGKLSVTDAAG 1349
QY 1627 NMVNVRRDSCMPLWLVVDPQVYVMTGNTSALKSVYTTQGHELAMMTYHNSGILLATKS 1686
DB 1350 NKVFLRDYTSQVNSIENTKQCKELRMTMKMLHELSTPDNYNVVYEHGPTGLLRKL 1409
QY 1687 NENGWTFYVDSFGRLTNVTFTPGQVSSFRSDTD-SSVHVQVETSCKDDVTITNLSAS 1745
DB 1410 DSTGRSYVNYNDEFGRLTSAVTPTRGVIELSFDLSVKGAVKVENAKEM-----SLLIQ 1465
QY 1746 GAFYTLLOQDVRSNYIYGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVG-----KR 1798
DB 1466 GATVIVRNGAASRTTYVMDGSTTSITPWGHNLOMEVAPYITLAEQOFLGESVVPVAK 1525
QY 1799 NVTLPIDNLGLNLEVR-----QRKEQARG---QVTFGRRLRVHNRNLLSLDFDRV 1846
DB 1526 RTEIAGDLA-NRFEMRYFVRQQLQAGKQSKGPPRPVTEVGRKLRVNGDNVLTLEYDRE 1584
QY 1847 TRTEKIYDHRKFTILILYDQAGPSLWSPSSRLNG-----VNVYSPGYYIAGIQGIMS 1902
DB 1585 TQSVVVMVDDKQELLNVTYDRTSRPISFRPOS---GDYAYVDLEVDYRGLRVSWKGVLIQ 1641

QY 1903 ERMEYDOAGRITSRIPADGKTMVYTYLEK--SMVLLHSORQ--YIFEDKNDRLSSVTM 1958
 Db 1642 EAYSFDRNGRLNEIKYGDGSTMVYAFKDFGSLPLKVTTPRRSDYLLQYDDAGALQSLLT 1701
 QY 1959 P--NVARQTLTETRSVGYRNIYQPPEGNASVQDFTEDGHLLHTFYLGTGRVYKYGK 2016
 Db 1702 PRGHIAFSLQ--SLGFFKYQVYSPINRHPPEILYNDEGQILAKIHPHQSGKFAVHDT 1759
 QY 2017 LSKLAETLYDTTKVSTYDETACMLKTNLQNEGPTCTTIRYQIGLIDRQIFRP--TEEG 2075
 Db 1760 AGLETLIAGLSHTHYQDTLSVKSVEVQEPFELRREFKYHAGILKDEKLKRGSKNS 1819
 QY 2076 MVNARPDYNDNFRVTSQAVINETPLPIDLYRYDDVSKTEQFQKFGVIVYDI--N 2131
 Db 1820 LASARYKAYDGNARLSGIEAIDDKELPTTRKYQNQLGLE-----VVQDLKITTEN 1872
 QY 2132 QIITAVMTTKHIF-----DAYGRKEVQYELFRSLMTVMVQYDNMRGVKKELKVP 2185
 Db 1873 AFNRTVIQDSAKQFAIVDYDQGRVKSVMNMYNIDVFRLELDYDLNRKIKSQKTFEG- 1931
 QY 2186 YANTRYSYDYDAGO--LOTYSINDKPLMRYSYDLNGN--LHLLSGNSARLTPLEYDID 2243
 Db 1932 -RSTADPKYNADGHVEVLGN--WKYLPDENGNVGVVDQGEKEN-----LGYDID 1984
 QY 2244 RITRLGVQY--KMDPEGLRQGRGDIPEYNSAGLLIKAYNRA--GWSVRYRYDGLGRV 2300
 Db 1985 RVIKGVGVEPNYDARGVFKRGQYKRYNNRQQLIHSPEERFQSW---YYDDRSRLV 2041
 QY 2301 SSKSSHHLQFFYADLTPTKTHLYNHSSEISLYYDLOGLHFAELSSGDSFFYATC 2360
 Db 2042 AWHDNKNTQYIYAPRPHLVTHVHPKIKRTMKLFYDDRLMLALE--HEDQYIYAT 2100
 QY 2361 DNTGTPLAVFSGTGLMIKQIYATAYGEIYMDNPNFQIIGYHGGIYDPLTKLHVHMRD 2420
 Db 2101 DQNGSPLAFDQNGSIVKEMKRTPFGRITKDKPEFFVDFHGGIYDPLTKLVYTEQ 2160
 QY 2421 YDVLGRWTSPPDELKWLHSSNMVFN--LYMFKNNPISQDIKPMTDVNSWLLTF 2478
 Db 2161 YDPHVQWMTF---LWETLATEMSHPTDFIYRYHNDPI--NPNKPNYMDLDSWLOJF 2216
 QY 2479 GFOLHNVIPGYPKPDMAPEPSVELIHTQMTQEMDNSK-----ILGVOCEYQKOLK 2531
 Db 2217 GYDLNN-----NQSRYTKLAQYTPQASIKSNLTAPDFGVISGLECIYEVK--- 2261
 QY 2532 AFTLBRPOLYSTTSCQAPKTKKFPASSG---VEGKGVKPAKDGRTVTDIISVA 2587
 Db 2262 ---TSEKFSDFDPVFKPLKTEPKMKNLLPRVSRGVEGEGVLLSRIGRGALVSVVDGS 2318
 QY 2588 N---EDGRVAATLNHAHYLENLHFTIDGVDTYFVKPSPSGDLAILGLSGRTELENG 2644
 Db 2319 NSVVQD--VVSFVNSYFLD--LHFSIHQDVFYVK---DNVLKRDNEELRLGGM 2371
 QY 2645 VNVTSQINTVLNGRTRRYTDIQLQXGALCLNTRYGTTLDDEKARVLELARAQVRAWA 2704
 Db 2372 FNISTEISDHGSAK--ELRLHGPDAWIYKGVDPQERHRLILKHAHRAVERAW 2428
 QY 2705 REOQRLREGEGLRAWTEGEKQOVLSTGRVQYDGFVLSVEQYELSDSANNHFMQS 2764
 Db 2429 LEKQVAAFGQGRGDWTEEEKSELVQHDVQDWMNGIDHSHIKYPOLADDPGNVAFQDA 2488
 QY 2765 EMGR 2769
 Db 2489 KRKR 2493

RESULT 3
 A54148
 cdz protein - fruit fly (Drosophila sp.)
 C:Species: Drosophila sp.
 C:Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 21-Jul-2003
 C:Accession: A54148
 R;Levine, A.; Bashan-Ahrend, A.; Budai-Hadrian, O.; Gartenberg, D.; Menasheerow, S.; Wide

Cell 77, 587-598, 1994
 A:Title: odd 02: a novel Drosophila pair rule gene
 A:Reference number: A54148; MUD:94243925; PMID:7514504
 A:Accession: A54148
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-2406 <LEV>
 C:Genetics:
 A:Gene: FlyBase:Ten-m
 A:Cross-references: FlyBase:FBgn0004449
 F:514-540/Domain: EGF homology <SGF>
 F:610-637/Domain: EGF homology <EGF1>
 Query Match 23.6%; Score 3515.5; DB 2; Length 2406;
 Best Local Similarity 34.0%; Pred. No. 5.7e-187;
 Matches 724; Conservative 433; Mismatches 943; Indels 130; Gaps 42;
 QY 384 EDTASSVPVTDVSLYPSGGTGLTPDRKKGTEKSPSPFPEDSFIDSGEIDVGRAS 443
 Db 347 EATSSAATSSSSQSLTSPSSSLANANNNGARTF---PARSPDPG-TTFQGITLQKLT 402
 QY 444 QKIPPGFWRSQVFDIDHPVHLKFNVSGLKAAALVIGRKLPPSHQTFQFVLLDGERLL 503
 Db 403 KEIQPSYWNMQFYQSEPAIVKFDYIIPRGASIGVGRNALFTHTQYHFKVLSFSAS 462
 QY 504 TQEARS--LEGTPRQSRGTVPSPSHETFTQYLDLSGWHIAFYN-DKSESVVSFLTTAI 560
 Db 463 TRTARAHLSTIREVTR-----YMEPGHWFVSLYNDGDDVQELTFYAAVAE 508
 QY 561 ESDVNCNSCYNGDCISGTCHGFLGPDGCRASCPVLCSNGOYMKGRCLCHSGMK 620
 Db 509 DMTQNCNCGSGNGQCLLGHCCQCPGFGDDCESVCPVLCSQHGHTYNGECICNFGWK 568
 QY 621 AEDVPTNQCIDVACSNHGTCTGTCTCINPGYKESCEEVDMDPTCSRGRGVCRGEC 680
 Db 569 KESLRHDECEVADCSGHGHCVSCKQCMRGYKGFCEEVDCHPHNCSGHGFCADGTC 628
 QY 681 FVGWGTGNCST---PRATCLDQCSGHGTFLPDGLCSDDPSWTGHDCSIEICAADCGHG 737
 Db 629 KKGWKGPDCAITMDQALQCLPDCSGHGTFLDPTQCTCEAKNSGDDCKELCDLDCGQ 688
 QY 738 VCVGTCRCRDGWMGAACDORACHPRCAEHGTCDKGCSCPGWNGEHTCTIAHYLDRVVK 797
 Db 689 RCEGDACAPCEWGEYCNTRLCVRCNEHGCQCKNGTCLCVGWNGKCTI----- 739
 QY 798 EGPGLCNGRCTLUNG--WHCVQOLGMRGACDTSMETACGDSKNDNDGDLVDCMDPD 856
 Db 740 EGPNSCAGHQCRVSGEGWECRCYEGMDPGDCGIALELNCGDSKNDKDLVDCEDPE 799
 QY 857 CCLQPLCHINPLGSGNPLDIIQETQVPVSQONLHSHFYDRIKFLVGRDSTHII PCENPF 916
 Db 800 CCASHVCKTSQLCVSAPKPIDVLLRKQPPAITA---SFFERMAFLIDESLQYAKLET 856
 QY 917 DGHACVIRQVMTSDGTPLVGVNISFVNNPLFGYITISQDGSFIDLVTNGGISIIILFER 976
 Db 857 NESRSVIRGRVVTSLGMLGVGRVS--TTLLSGFTLRDDGWFDLWMVGGGAVTLQFGR 915
 QY 977 APITQHTLWLWDWFFVYMETIIMRHENE-----IPSCDLSNFARNPVPVSPFLS 1030
 Db 916 APRFQSRIVQVFWNEVVIIDLVMSNBEKGLAVTTTTCFAHDYDLMKVPVLAGWK 975
 QY 1031 FASSCAEKGPVPEIQALQEEISISGCKMRLSYLSRTPGYKSVLRISLTHPTIPENLMK 1090
 Db 976 FQACPDPSAILAESQVIOESLQIPGTGLNLVYHSSRAAGYLSTIKLQTPDVPVPSLH 1035
 QY 1091 VLMVAVEGLFRKWFAPADLSYYFINWKTVDVYNOKVFGLSBAFVSQVYEVESCDLIL 1150
 Db 1036 IHLRTIEGLIFRIFEADFGIKFTYAWNRLNRYORVYGVTTAVVKVGYQYVTDCTD-IV 1094
 QY 1151 WEKRTVLOQYEDASKLGGWSDKHALNIQSGILHKGNGENQFVSQQPPVIGSMNG 1210
 Db 1095 WDIQTTKLSGHDMSISEVGVGNLIDIHRYNFHIGILQKGGSNILYLRNKRPIILLTMGD 1154

Db 417 HVSLLAEEAEGVSTTCPCNDSCGRGSCYLKCDIDGYQGVDCSKVCPVLCSAHHYGG 476
Qy 610 GRCLCHSGWGAECDFVPTNOCIDVACSNHGTCTGTCTICNPGYKGSCEVDCMDPTCSG 669
Db 477 GVCHCEGWGAECDFVGECEVFNCSHGRCLEGECHERGWKGYCQDHCLDPLCSG 536
Qy 670 RGVCVRGECHCFVWGNTCET---PRATCLDQCSGHGTFPLDTGICSDPSPWTGDCSI 726
Db 537 HGTCAQCQCKAGWOGEDGCTTQQVYQCLPGLCSBHGTYDLETGQVCERHWTGDCSQ 596
Qy 727 EICAAOOGGHVCGVGTCTCEDGMMGAACDORACHPRCAEHGTCRDKGCECSPGNGBHC 786
Db 597 AVCLSDCGRNGVCSHGKRCNSGWTGMLCDQLPCDSRCSHGCKGTCTVCSQGWNGRHC 656
Qy 787 TIAHYLDVRVKEGCPGLGNGRCRTLDLNGHCVCLGWRGAGCDTSMETACGDSKDNDS 846
Db 657 TL-----PCENGCSRHGQCTLENGEYRCDCTEGWAGSDCSIALELNCKDNIDNG 707
Qy 847 DGLVDCMDPCCLOPCHINPLCLGSPNPLDIQTQVPSQQNLHSFYDRKFL 901
Db 708 DGMTDCSDSCSHAPACSEHIMCLSSNDPVEVLLRKQPP---SVTASFYQRYKFL 759
RESULT 7
T12457
hypothetical protein DKFP56400423.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 23-Jul-1999
C;Accession: T12457
R;Wambutt, R.; Heubner, D.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z17524
A;Accession: T12457
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-184 <WAM>
A;Cross-references: EMBL:AL080120
A;Experimental source: fetal brain; clone DKFP56400423
A;Genetics:
A;Note: DKFP56400423.1
Query Match 6.3%; Score 943; DB 2; Length 184;
Best Local Similarity 99.5%; Pred. No. 1.8e-45;
Matches 182; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Qy 2587 ANEDGRVAAILNHAHYLENLHFTIDGVTHYFVKPSPSEGLAILGLSGGRTLENGVN 2646
Db 2 ANEDGRVAAILDHAHYLENLHFTIDGVTHYFVKPSPSEGLAILGLSGGRTLENGVN 61
Qy 2647 VTVSQINTVLNGTRTRYTDIQLQYGCALCNTRYGTTLDDEKARVLELAPQRAVQAWARE 2706
Db 62 VTVSQINTVLNGTRTRYTDIQLQYGCALCNTRYGTTLDDEKARVLELAPQRAVQAWARE 121
Qy 2707 QQRLEGEGLRAWTEGEGQQVLSTGRVQYGGFFVISVEQYPELSDSANNHFRQSEM 2766
Db 122 QQRLEGEGLRAWTEGEGQQVLSTGRVQYGGFFVISVEQYPELSDSANNHFRQSEM 181
Qy 2767 GRR 2769
Db 182 GRR 184

RESULT 8

JQ1322
tenascin precursor - mouse
N;Alternate names: contactin; hexabrachion
C;Species: Mus musculus (house mouse)
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000.
C;Accession: JQ1322; A37936; B37936; S14571; S50209
R;Saga, Y.; Tsukamoto, T.; Jing, N.; Kusakabe, M.; Sakakura, T.
Gene 104, 177-185, 1991
A;Title: Murine tenascin: cDNA cloning, structure and temporal expression of isoforms.

A;Reference number: JQ1322; MUID:92009211; PMID:1717349
A;Accession: JQ1322
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-2019 <SAG>
A;Cross-references: GB:D90343; NID:9220609; PIDN:BAAL4355.1; PID:G220610
A;Experimental source: cell line 2H6GR
A;Note: the authors translated the codon ATG for residue 60 as Trp
R;Weller, A.; Beck, S.; Ekblom, P.
J. Cell Biol. 112, 355-362, 1991
A;Title: Amino acid sequence of mouse tenascin and differential expression of two tenascin
A;Reference number: A37936; MUID:91107734; PMID:1703162
A;Accession: A37936
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-201, 'E', 203-317, 'S', 319-620, 622-1010, 'N', 1012-1018, 'S', 1020-1024, 'H', 1026-
A;Cross-references: GB:X56304
A;Accession: B37936
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-201, 'E', 203-317, 'S', 319-620, 622-1010, 'N', 1012-1018, 'S', 1020-1024, 'H', 1026-
A;Cross-references: GB:X56304
R;Weller, A.; Beck, S.; Ekblom, P.
submitted to the EMBL Data Library, August 1990
A;Description: Amino acid sequence of mouse tenascin and differential expression of two
A;Reference number: S14571
A;Accession: S14571
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-201, 'E', 203-317, 'S', 319-1018, 'S', 1020-1024, 'H', 1026-1305, 'S', 1307-2019 <WE
A;Cross-references: EMBL:X56304; NID:G54768; PIDN:CAA39751.1; PID:G54769
R;Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A;Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues by
A;Reference number: S50206; MUID:95035091; PMID:7524681
A;Accession: S50209
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: mRNA
A;Residues: 46-146 <GUJ>
A;Cross-references: EMBL:X80281
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994
C;Superfamily: tenascin; EGF homology; fibrinogen beta/gamma homology; fibronectin type
C;Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extracell
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-2019/Product: tenascin, long splice form #status predicted <MAT>
F;23-1071, 1527-2019/Product: tenascin, short splice form #status predicted <MAT2>
F;408-434/Domain: EGF homology <EGF>
F;622-703/Domain: fibronectin type III repeat homology <FN3A>
F;711-793/Domain: fibronectin type III repeat homology <FN3B>
F;802-884/Domain: fibronectin type III repeat homology <FN3C>
F;892-976/Domain: fibronectin type III repeat homology <FN3D>
F;984-1064/Domain: fibronectin type III repeat homology <FN3E>
F;1073-1155/Domain: fibronectin type III repeat homology <FN3F>
F;1346-1428/Domain: fibronectin type III repeat homology <FN3G>
F;1437-1519/Domain: fibronectin type III repeat homology <FN3H>
F;1527-1608/Domain: fibronectin type III repeat homology <FN3I>
F;1616-1696/Domain: fibronectin type III repeat homology <FN3J>
F;1704-1784/Domain: fibronectin type III repeat homology <FN3K>
F;1799-2007/Domain: fibronogen beta/gamma homology <FBG>
F;38, 166, 184, 327, 788, 1018, 1079, 1093, 1119, 1184, 1210, 1275, 1301, 1354, 1364, 1394, 1627, 1878, 19

Query Match 4.4%; Score 659.5; DB 1; Length 2019;
Best Local Similarity 19.5%; Pred. No. 5.4e-28;
Matches 407; Conservative 273; Mismatches 714; Indels 689; Gaps 97;

Qy 566 CPSCYCNCGPCISGTCHCFGLGPGCGRASCVPVLCSGNGCYMKGRCLCHSGWKGAEC-- 623
Db 315 CPNDFRGRINGTCYCEGFTGEGELTCENDQGRGCGEGQVCNEGFPAGADCSSE 374
Qy 624 -----DVPTNQCIDVACSNHGTCTGTCTICNPNPYKGS 656
Db 375 KRCPADCHRRGCLNGQCECDGFTGADCDGLQCPN-GCSHGRCVNGQVCDCGYGTGD 433

QY 657 CEEVDCMDPTCSGRGVYRGECHCFVWGNTNCETPRATCLDCSCGHGTFPLDTCGLCSCD 716
 Db 434 CSORRCFN--DCHNRGLCVQKCI--CEQFNGFDCV--EMSCPNDCHQGRGV--NGMCICD 488
 QY 717 PSMTGDCSIEICAAACGGHGVCGGTCTRCEDGWMGAACDQACHPRCAEHGTCRDGKCE 776
 Db 489 DDYTGDCDRRCPRDCSQGRGVCGOCICEDGFTGPDCAELSCPSDCHGHGRCVNGQCI 548
 QY 777 CSFGWNGEHTIAHYLDVVYKCECPGLCNGNHCETDLNNGHVCVCLGWRGAGCDSMET 836
 Db 549 CHEGFTGKCK--EQRCPDCHGQGR--EDGQCICHEGFTGLDCG--QR 592
 QY 837 ACSDSDNDGDGLVDCMDPC-CLQPLCHINPLCLGSPNPLDIIQETQVPSQNLHSPY 895
 Db 593 SCFDCSNQ--QCVSGRGICNBYGTGID--CSEVSPKOLI--VTEVTEETVNLAW 643
 QY 896 DR-----IKFLVGRDSTH-----IPGNPDGHCACVIRQVMTSDGTPLVGNV--- 940
 Db 644 DNMVRVTEYIMVTPHADGLEMQFRVPGD-----QTSTTIRELEPGVEYFI 690
 QY 941 --ISFVNNPLFGYTI--SQDGSPLVINGGSIILRPERAPFTQEHTLPLWDRPPVMT 998
 Db 691 RVFAILENKR-SIPVSARVATYLPAPEG-----LKFSIK-ETSVEVEWDPLDIAPETWE 743
 QY 999 IIMRHEENELPSCDLSNFARPWPSPPLTSPASSCAEKGPVPIQALOEIEISGCK 1058
 Db 744 IIFRNMNKEDEGEITKSLRPE-----TSYRQGLAPG-----QYEBIS--- 782
 QY 1059 MRLSYLSRTPGKSVLRISLTHTPTTFNLMKVLHV--AVEGRLPKFWFAAPDLSYF 1116
 Db 783 LHTVKNNTGFGKLVKVTTLRLDAPS-----HIEVKDVTDTALTITWFKPLAEID--- 831
 QY 1117 IWKTDVYNGKVGLEAFVSGYEVESCPDLILWEKRTVTLGQYIDASKLGGWSLDKH 1176
 Db 832 -----SLELSYGKDVPG-----DRTTIDLTEDNQYSIGNLRPDTE 868
 QY 1177 HALNIOGILHKGNGENQFVQOPVIGSIMNGRRSI--SCPSCNGLADGNKLLAPVAL 1235
 Db 869 YEVSLSIS-----RRVDMASNPAKETFTIG--LDAP--- 896
 QY 1236 TCGSDGSLVGVDFNYIRIPPSGNVTNILELRNKDPRHSHSPAHKYLATDPMGAVFLS 1295
 Db 897 -----RNLRVSQTNSTL-----EWRNVKADISYRIKAPISGG--- 933
 QY 1296 DSNRRRVFKIKSTVVVKDLVKNSEVAVAGTGDQCLPFDTRCGDGGKATEATLTPRGITV 1355
 Db 934 -----DHAEDVPSKQQA-----TTKTLTGLRPGI- 959
 QY 1356 DKFLIYFVDTMIRIDQNGIISTLLGNDLTSARPSCDSVMDISQVRLE--WPTDLA 1413
 Db 960 -EYGI-----GVSAYGDKESDPATINAATEIDAPKDL---RVSETTQPSLTFFFTPLA 1010
 QY 1414 -----IN-----PMDNGLYVLDNNVWLQISENHQVRIVAGRPMHC----- 1448
 Db 1011 KFDRLNLYSLPTGQSMEVQLPKDATSHVLT--LEPGQYTVLLIAEGRHKSPARV 1067
 QY 1449 -----QVPGTDHFLSKVA-----IHATLESATLAVSHN-----GVLY 1482
 Db 1068 KASTEVEPSLENLTVEAGWDLRLNLTADDLAYEYFVIQVQZANNVETAHNPFTVGNLR 1127
 QY 1483 IATDEKKINRIQVTTSGEISLVAGAPSCDCKNDANCDPFGSDGY-----AKDAXL 1536
 Db 1128 AADIPGLK-----VATSVRSYIGVA-----RGYRTPVLSAETGT 1163
 QY 1537 NTPSSLAVCADGELYVADLGNIRIFIRKNKPFNTQNMVELSPIDQELYLFDITGKHL 1596
 Db 1164 TTNL-----GEVTVAEVG-----WDAL---T 1182
 QY 1597 YTSQSLPTGDLVNYFTYTGDEGDIILTDNNGMNVNRDSTGMPJLWVDPG----- 1647
 Db 1183 LANWTAPEGAYKNFPIQVLEADTTQTQVN-----ITVPGGERSVDLPGL 1225
 QY 1648 ---QVYVWTM-----GTNSALKSVTQGHELAMMTYHGNSGLLATKSNENG---WTT 1693

Db 1226 KAAATRYITLKGVTQDFGT--APLSVEVLIEDLQOL-----GGLSVTEVSDGTLNWT 1278
 QY 1694 FPEYDSFGLRNLVFTPTGQVSSPFRDSSVHVQVETSSKDDVTITTNLSASGAFYTL-- 1751
 Db 1279 -----DOLAYKHVVQVQZANNVBAQAQNLTVFVPSGLRAVDI 1313
 QY 1752 ---LQOVNRNYYIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVKGKENVTLPIDNG 1807
 Db 1314 PGLKADTPYRVSIIYGVTOGYTPLMLSDTVSTARE-----PEIGNLVNSDVTPKS 1362
 QY 1808 LNLVWQRQKEQARGQVTVFGRRLRVHNRNLLSDFDRVTRTEKIYDDHRRKFTLRILYDQ 1867
 Db 1363 FNL--SW-----TATDGFID---MFTIEII--- 1382
 QY 1868 AGRSLWSPSRNLGNVNTVSPGG-----YIAGIQGIMSERM-----EYDQA 1910
 Db 1383 DSNRLQQAENISGAERTAHISGLPPSTDPIVLSGIAPSIKTKTSTTATTAEPEVD 1442
 QY 1911 GRISRIFADKGTWSTYTL--KSMVLLHSQOYIFEFKNDRLSSVTMPNVAR--QTL 1966
 Db 1443 NLLVSDATPDGFRLSWTADGIFDSFVIRIDTK-----KQSEPOEISLSPERTDI 1495
 QY 1967 ETIRSVGYR-NIYOPPEGNASVIQDFTEDOHLLHTFVLTGRRVVIYKGLSKLAETL- 2024
 Db 1496 TGLREATEYEIELYCISRRRS-----QPVSALATAMGSPKEIM--FSDITENAAATVS 1547
 QY 2025 --YDTTKVS---FTY-DETAGMLKTLNQNEGFTCTIRYROIPLIDROIF-----RPT 2073
 Db 1548 WRAPTAQVESPRIYVPMTGAPSMVTV--DGTOTETRLVLTGVEYRVSIVIAMKGEE 1605
 QY 2074 EGMVNARPDYNDNS-----FRVTSMAVINETP--LPIDLYRYD-----DVS 2115
 Db 1606 SDPVSGTLITALDGPSSGLLIANITDSEALAMWQPAIATVDSYVTSYTGVRVPEVTRTVSG 1665
 QY 2116 KTEQPGKEGV-----IYDINOIITAVMT-HKHPDAYGRMKEVOYEIERSLMY 2165
 Db 1666 NTVEYHDLSPATEYILSIFAEKQOQSSTIAKFTTDLDSPREFTATEYQSETALLTW 1725
 QY 2166 MTQVDNMGVRVVKELKVPVANTRYSEYDA-DGOLQTVSINDKPLWRSYDINGNLH 2224
 Db 1726 RP-----PRASVTGYLLVSVSDGTVKEVIGVDPDTT---SYSLAD--- 1762
 QY 2225 LLSPON--SARLPLRVDIRI-----TRLGVOYKMDGDFLRQRGDIFEVNSAGL- 2276
 Db 1763 -LSPSTHYSARIQALSGSLRSKLITQITFTTIG-LLYPPPRDCSOAMLNGD---TTSGLY 1816
 QY 2277 -----LIKAV---NRAGSVRYRYDGG-GRVSSKSSSHHLQF-----FVAD 2316
 Db 1817 TYINGDKTQALEVYCDMTSGGWIPLRRKN-CREDFYRNWKAAYAAGFDRREFFWIG 1875
 QY 2317 LTNPTKVTHLYNHSSEITSYLDLQGHIFAMELSSGDEFYIACDNIGTPLAVFSGTGLM 2376
 Db 1876 LDNLSKITAOQY-----ELRVLDQD-----GESAVAVYDRPSVG----- 1911
 QY 2377 IKQILYTAGIYVDTNPNFQI-IIGHYHGLYDPLTKLVHNGR 2418
 Db 1912 -----DAKSRYLKVEGYSGTAGDSMN--YHNGR 1938

RESULT 9

S19694
 tenascin precursor - pig
 N:Alternate names: contactin; hexabrachion
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: S19694
 R:Nishi, T.; Weinstein, J.; Gillespie, W.M.; Paulson, J.C.
 Eur. J. Biochem. 202; 643-648, 1991
 A:Title: Complete primary structure of porcine tenascin. Detection of tenascin transcri
 A:Reference number: S19694; MUID:92104189; PMID:1722152
 A:Accession: S19694
 A:Molecule type: mRNA

A;Residues: 1-1746 <NIS>
A;Cross-references: EMBL; X61599; NID:G2124; PID:CAA43796.1; PID:G2125
C;Superfamily: tenascin; EGF homology; fibrinogen beta/gamma homology; fibronectin type
C;Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extracellular
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-1746/Product: tenascin #status predicted <MAT>
F;346-372/Domain: EGF homology <EGF>
F;377-403/Domain: EGF homology <EGF>
F;622-703/Domain: fibronectin type III repeat homology <FN3A>
F;711-793/Domain: fibronectin type III repeat homology <FN3B>
F;802-884/Domain: fibronectin type III repeat homology <FN3C>
F;882-976/Domain: fibronectin type III repeat homology <FN3D>
F;984-1064/Domain: fibronectin type III repeat homology <FN3E>
F;1073-1155/Domain: fibronectin type III repeat homology <FN3F>
F;1164-1246/Domain: fibronectin type III repeat homology <FN3G>
F;1254-1335/Domain: fibronectin type III repeat homology <FN3H>
F;1343-1423/Domain: fibronectin type III repeat homology <FN3I>
F;1431-1511/Domain: fibronectin type III repeat homology <FN3J>
F;1526-1734/Domain: fibronectin type III repeat homology <FN3K>
F;38,166,184,327,788,1034,1079,1121,1354/Binding site: carbohydrate (asn) (covalent) #sb

Query Match 4.4%; Score 652; DB 1; Length 1746;
Best Local Similarity 21.3%; Pred. No. 1.1e-27;
Matches 342; Conservative 204; Mismatches 558; Indels 512; Gaps 72;

QY	566	CPNCGVNGDCISGTC	CHFLGFLGPGCGR	ASCPVLCSGNGQVWKGRCLCHSGWKAEC	624
DB	315	CPKDCDFRGCINGTC	YCDGFEDEGCGRL	ACPHGCRGRCGEGQVCDGEGAGACSE	374
QY	625	--VPTNOCIDVACN	GTCTGTCICNPGV	KGSCREVCMDPTCSGRGVGKSGCHFCV	682
DB	375	RCPCSD	-----CHNRC	LDRCRCDDGFEGEDCCELR	427
QY	683	GWGTCNCTPRATC	DOCSGHGFTPLT	GLCSDPSWTGHDSCIECAADCGHGVCGV	742
DB	428	GRTEGDCSOLR	--CPNDC	CHGRGCV--QGRCEHGFQYDGCSEMSCPDCHQHGR	483
QY	743	TCRCEDGMGAAC	DQACRACHTCR	GDGKCEGSPGNGBHCITIAHLVDRVKGCGP	802
DB	484	MCVDDGTGDCR	ELRCPGDCSQR	CGVDRGRCVCHGFGAGPDC-----ADLACPS	534
QY	803	LCNNGRCGLD	LNWHCVCOQ	HRGAGC-----DTSMETACGDSK--DNDGDGLVDCM	853
DB	535	DCHGRGRC	--VNG-QVC	HEGFTGDKQRCRCPGDCGQRCVDCQVCHEGFTGLD	590
QY	854	DPDC	-----CLQPL	CHINPLCLG-----SNPLDIIQETQVPSQQLHSPYDR	897
DB	591	QRSCPND	CNMGQVSGRC	CNMGYSGEDCSQVSP-PKDLI--VTEVTEETVNLWDNE	646
QY	898	---IKFLVGRD	STH-----IIPEN	PFDDGHACVIRGQVMTSDGTPLVGN-----I	941
DB	647	MRVTEVLIV	TPTHEDGLEQ	RVFDC-----QSTTIRELEPCGVFIRVF	693
QY	942	SFVNNPLFG	VTISRQDGS	DLVTNGGSIILRPERAPFTQSHTLWLWDRFPVNETIM	1001
DB	694	AILENKK-SIP	VSARVATY-LTP	PEG-----LKFPSIK-BTSVEVEMDPLDIAFETWEIIF	746
QY	1002	RHEENELPS	CDLNFARPNV	PSPLTSPASSCAEKGPVIEIQALOEIISICKMRL	1061
DB	747	RNNKDEGE	EITKSLRPE-----	TYRQTGLAPG-----QEVEIS--LHI	785
QY	1062	SYLSSRTPG	YKSVLRISL	THPTIPFLMKVHLMAVVEGLFRKWFAPADLSYFIWDT	1121
DB	786	VKNTRGPG	LKRVTTTRLD	APS-----QIEAKDVTDTTALITWFKPLAEID-----	831
QY	1122	DVYNQV	FGLSEAFVSG	VEYESCDPLILWEXRTVLQGYEIDASKLGWSLDKHALNI	1181
DB	832	-----GIB	LTGKDVPG-----	DRTTIDLTHEENQVSGISGNLKPDTTBYEVS	873
QY	1182	QSGILK	KGNGENQFV	SQQPFVSGINGNRRRSISCPSCNGLADGNKLLAPVALTCGSDG	1241
DB	874	IS	-----RR-----	ADWSSNPAKETFTTGLDA	895

QY	1242	SLYVGDFNYIRRI	PFPSGNVTNILELR	NKDFRHSHPAKHYLATDPMGSAVFLSDSNRR	1301
DB	896	P-----RN	LRIISQD	NSITL-----EWENKAAADTYRIKAPISGG-----	933
QY	1302	VFKIKSTVVVK	LDVKNSEVAGT	GDQCLPFDDTRCGDGGKATATLTNPGRITVDKFLGI	1361
DB	934	-----DHA	EVPRSPQ-----	TTTATLTGLRPGT--EYGI-963	
QY	1362	YFVDGTMIRRI	DONGIISTLLGS	NDLTSARPLSCDSVMDISQVRLEWPTDLA-----IN	1415
DB	964	---GVS	AKDKESD	PATINAATDLPKDFRVSLEKE--SSITLLWRTFLAKEDRYRLN	1018
QY	1416	-----PMD	SLVLDNNVYLQ	ISENHQRVIRVAGRMHCQVPG-----1452	
DB	1019	YGLPSGP	VEVQLPRNAT	SVILRG---LEPGQEYITLLTAEGRHKSKPARVKASTAGEP	1075
QY	1453	-----IDH	FLSKVAIHAT	LESATALAVSHNGVLYIAETDEKINRI-RQVTT	1500
DB	1076	EIGNLS	VDITPES	FSLSWTATEGAFETFIETIDSNRFL---ETMEYNISGAERTAHIS	1132
QY	1501	G-----EIS	VAGAPSGC	CDKNDANCDFCSDGDDGYAKADAKLNTPSLSLAYCADGE-----1549	
DB	1133	GLRPGND	FIVYLSGL	APGIQTK-----PISATATTEAEPEVDN	1170
QY	1550	LYVADL	GNIRIRFK	RKPKPFLNTQNMVELSSPIDQELY-----LPDITCKHLYTQSLPT	1603
DB	1171	LLVSDA	-----TPD	FLSWTADGVDGFSVLKIRDTK-----KQSEPL	1209
QY	1604	GDLYNFTY	GDGDTILIT	DNNGMNVNRDSTGMPLWLVPDGVVWVTMTGNSALKSV	1663
DB	1210	-----EIT	LASERTED	IGLEAEYEI-----ELYGISGKES-----1244	
QY	1664	TTQGH	ELAMTYHNS	GLLATKSNENGWTFYEDYSGRLTNVTFPTQGVSSPRTDSS	1723
DB	1245	-----QPV	SAIATAMG	SPKEITFSDITENSATVSMWVPTAQVESFR-----I	1287
QY	1724	VHVOVETSS	KDDVTIT	TNLSAGAFYTLQDVNRNSYIGADGSLRLLLANGMEVALOTE	1783
DB	1288	TYVPTG	ASVTVT	VDGKTQTRLLRL-----PGVEYLVSVIAVKGFE---ESE	1334
QY	1784	PHLAGT	VNPVTKRNV	TLFDNGLNLVWQRKEQARGQVTFGRRLRVHNRNLLSLOF	1843
DB	1335	P--VSGT	-----LT	ALDGPGLVT-----1352	
QY	1844	DRVTRTEKI	YDDHKKF	TLRLYDQAGPSPWSPS--SRNLGNVNTYSPGGYIAGIQRGIMS	1902
DB	1353	ANITDSEAL	-----AMQ	PAIAPVDHVISYT-GRVPEITRTVSG	1392
QY	1903	ERMEY	-----DQ	AGRITSRIFAD-GKTWSYTYLEKSMVLLHSQRYIFEFDKXNDRLSSV	1956
DB	1393	NTVEYAL	TNLEPATE	YTLRFAEKGPKQKSTITTK-----FTTDL-DSPRDL	1438
QY	1957	TMNVARQ	T-ETIR	-----SVGYRNIYQPPEGNASVI-----QDFTEDGHLLHTFYLG	2005
DB	1439	TATEVQ	SETALLTW	PPRASVTGYLLIYVESVDGTLKEVWVGPETTSYLSGLSPSTHYTA	1498
QY	2006	TGRVIVKY	GKL-SK	LAETLYDTTKVSFTY-----DETAGMLKTNLONE	2049
DB	1499	---RIQ	ALNGFLSK	SKSQVTFITGLLYPPRDCSQAMLNGDTTSG-LTYIYNND	1550

RESULT 10
A32160
tenascin-C - human
N:Alternate names: hexabrachion
C:Species: Homo sapiens (man)
C>Date: 31-Jul-1989 #sequence revision 12-Apr-1996 #text change 20-Aug-1999
R:Accession: I38337; A32160; S16166; S50208; S49354
R:Gherzi, R.; Carnemolla, B.; Siri, A.; Ponassi, M.; Balza, E.; Zardi, L.
J. Biol. Chem. 270, 3429-3434, 1995
A:Title: Human tenascin gene. Structure of the 5'-region, identification, and characteri
A:Reference number: A55974; MUID:95155442; PMID:7531707
A:Accession: I38337

[illegible]

Qy 1661 K-SVTTQGHAMWYHNGSGLLTKSNENGWTFYVEYDFGLRNTVTPFTQGVSRSD 1719
Db 1230 HYITIRGVQDFSTFSLSEVLETPDMGNLTVEVSDALRLNWTTPDGYDQFT-- 1287
Qy 1720 TDSVHVQVETSSKDDVTIITNLSASGAFYLLQDQVRNSYIIGADGSLRLLANGMEVA 1779
Db 1288 -----IQVQEA-----DQVEAHNLTVPGLSR-----SMEI- 1313
Qy 1780 LQTEPHLLAGT 1790
Db 1314 ----PGLRAGT 1320

RESULT 11
A32230
N;Alternate names: cytotactin; hexabrachion
N;Contains: tenascin 190K; tenascin 200K
C;Species: Gallus gallus (chicken)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999
C;Accession: A32230; B32230; A33379; B33379; S01292; A30903
R;Jones, F.S.; Hoffman, S.; Cunningham, B.A.; Edelman, G.M.
Proc. Natl. Acad. Sci. U.S.A. 86, 1905-1909, 1989
A;Title: A detailed structural model of cytotactin: protein homologies, alternative RNA
A;Reference number: A32230; MUID:89184536; PMID:2467292
A;Accession: A32230
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1810 <JON>
A;Cross-references: GB:J04519; NID:G211717; PID:AAA48745.1; PID:G211718
A;Accession: B32230
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1044,1318-1810 <J02>
A;Cross-references: GB:J04519
R;Spring, J.; Beck, K.; Chiquet-Ehrismann, R.
Cell 59, 325-334, 1989
A;Title: Two contrary functions of tenascin: dissection of the active sites by recombina
A;Reference number: A33379; MUID:90030407; PMID:2478295
A;Accession: A33379
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-204, 'G', 206-221, 'A', 223-380, 'D', 382-386, 'H', 388-444, 'HN', 447-450, 'V', 452-5
A;Cross-references: GB:M23121; NID:G212746; PID:AAA49086.1; PID:G212749
A;Accession: B33379
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: mRNA
A;Residues: 1-204, 'G', 206-221, 'A', 223-380, 'D', 382-386, 'H', 388-444, 'HN', 447-450, 'V', 452-5
A;Accession: C33379
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: mRNA
A;Residues: 1-204, 'G', 206-221, 'A', 223-380, 'D', 382-386, 'H', 388-444, 'HN', 447-450, 'V', 452-5
A;Cross-references: GB:M23121
R;Pearson, C.A.; Pearson, D.; Shibahara, S.; Hofsteenge, J.; Chiquet-Ehrismann, R.
EMBO J. 7, 2977-2982, 1988
A;Title: Tenascin: cDNA cloning and induction by TGF-beta.
A;Reference number: S01292; MUID:89030589; PMID:2460335
A;Accession: S01292
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 27-181, 'R', 183-204, 'G', 206-221, 'A', 223-380, 'D', 382-386, 'H', 388-444, 'HN', 447-
A;Cross-references: EMBL:X08030
C;Note: part of this sequence was confirmed by protein sequencing
C;Superfamily: tenascin; EGF homology; fibronogen beta/gamma homology; fibronectin type
C;Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extracell
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-33/Domain: propeptide #status predicted <PRO>
F;34-1810/Product: tenascin 230k #status predicted <MAT>
F;223-249/Domain: EGF homology <EGF1>
F;316-342/Domain: EGF homology <EGF>
F;592-673/Domain: fibronectin type III repeat homology <FN3A>
F;681-765/Domain: fibronectin type III repeat homology <FN3B>
F;773-857/Domain: fibronectin type III repeat homology <FN3C>

F;865-949/Domain: fibronectin type III repeat homology <FN3D>
F;957-1037/Domain: fibronectin type III repeat homology <FN3E>
F;1046-1128/Domain: fibronectin type III repeat homology <FN3F>
F;1137-1219/Domain: fibronectin type III repeat homology <FN3G>
F;1228-1310/Domain: fibronectin type III repeat homology <FN3H>
F;1318-1399/Domain: fibronectin type III repeat homology <FN3I>
F;1407-1487/Domain: fibronectin type III repeat homology <FN3J>
F;1495-1575/Domain: fibronectin type III repeat homology <FN3K>
F;1590-1798/Domain: fibronogen beta/gamma homology <FBG>
F;1734-1747/Domain: calcium binding #status predicted <CAB>

Query Match 4.3%; Score 641; DB 1; Length 1810;
Best Local Similarity 20.4%; Pred. No. 4.9e-27;
Matches 333; Conservative 219; Mismatches 563; Indels 520; Gaps 72;

Qy 543 FYNDGKESEVVSFTTAAIESVDNCPSCYNGYDNCISGTCCHFCFLGFLGPDGGRASCPVLCS 602
Db 332 FCEBSYTGEDCGELT-----CPNNGNGRCENGLCVCHGEFGVGDGDSQKRCPTCN 383
Qy 603 GNGQYMKGRCLCHSGWGAECB---VFTNOCIDVACSNHGTCTTGTCICINPGYKGSCEE 659
Db 384 NRGRCVDGRCVCHGEYLGEDCGELRCND-----CHNRGRINGQCVCDGEFIGSDCGE 437
Qy 660 VDCMDPTCSGRGVCVRGECFCVWGCTNCETPRATCLDCQSGHGTFLPDTGLSCDPSW 719
Db 438 LRCFN-DCQGRGRCINGQCEHGFGECDGELR--CPNDCNSHGRCV--NGQCVCDGEY 492
Qy 720 TGHDCSIEICAADCGHGVCGVGTCTCEDGWMGAACDQACHPRCABHGTCTRGKCECSP 779
Db 493 TGEDCGELRCFNDCHNRGRCVGRGVCVCDNGFMGDCGELSCPNDCQHGRCVDRGRCVCH 552
Qy 780 GWNGEHTIAHYLDVRVKEGCGPLGNGRCCTLDLNGHWCVCOLGWEGACSDTSMETACG 839
Db 553 GFTGEDCR-----ERSCPNDCNNVGR-----VEG-RCVCEGNG----- 587
Qy 840 DSKONDGDLVDCMDPDCCLQPLCHINPLCLGSPNPLDIIQETQVPVSOQNLHSFYDRI- 898
Db 588 -----IDCS-----VSP-----PTGLTVTNVTDKV--NLEWKHENLV 619
Qy 899 -KFLV-----GRDSTHIIPE-----NPDGGHACVIRQVMTSDGTPLVGVNISF 943
Db 620 NEYLTVVPTSSGGDLQFTVPGNQTSATHELEPGVEYFIRVAILKNKKS----- 672
Qy 944 VNNPLFGVTISRQDGSPLVTNGGISILRERAPFTQEBTLMWPWDRF---FVMEI 1000
Db 673 -----PVSARVATYLPAGE-----LFEKS-----VRETSVQVEWDPLISFDGWELV 715
Qy 1001 MRHEENIPCDL-SNFARNPVVSPPLTSPASSCAEKGPVPEIQALOEIISIGCKM 1059
Db 716 FRNMQKDDNGDITSLKRPE-----TSYMQPGLAPG-----QYNNVS---L 754
Qy 1060 RLSVLSRTPCYKSVLRISLTHPTIPNLKMKVLMVAVEGLFRKWFAPADLSYFIWD 1119
Db 755 HIVKNTIRGELSRVITTKDAPS-----QIEAK-----D 784
Qy 1120 KTDVYNQKVFGLSEAFVSGYGEYSCPDLLWEXRTTVLQYE-----IDA 1165
Db 785 VTDT-----TALITWSKPLAEIEGIELTYGPKDVPDGRITDL 822
Qy 1166 SKLGWSLDKXHALNIOSGILHKNGENQVSOQPPVIGSIMNGRERS-----ISCPSC 1220
Db 823 SE-----DNQY-----SIGNLPHTEYEVEYTLIS 849
Qy 1221 NGLADGNKLLAPVALTCGDSGLYVGFNFYIRIFFSGNVNILELRNKDFRSHSFAHK 1280
Db 850 RGDMSDPA-----KEVFTDLDAPNLKRVSQTDNSITL---EWFESHANIDN 895
Qy 1281 YLATDPMSCAVFLSDNSNRVFKISVTVVKLVKNSEVAVAGTGQCLPFDDTRCDGGS 1340
Db 896 YRIKFAPIGG-----DHTELTVPKGNQA----- 919
Qy 1341 KATEATLTNPRGITVDKFGLIYFVDGTMIRRIDQNGIISTLLGNSDLTSARPLSCDSVMD 1400

Db 920 -TTTATLTGLRP-----GTEYGIGVAVRQ-DRESAPATINAGTDLNPKDLEVS DPT 971
Qy 1401 ISQVRLSWPTDLA-----INPMDSYVLDDNNVVIQISENHQVRIV 1441
Db 972 -TTLSLAWRPVAKFDRYLTVYSPSGKNEMIPVDSTFILRG---LDAGTEYITSLV 1027
Qy 1442 AGRPMHCQVP-----GIDHFLLSKVATHTALESATALAVSHNGV 1480
Db 1028 ABKORHKS KTTTGTSTBEPENLSVSTGWDGFGQLTWTADGAYE-----NFV 1078
Qy 1481 LYIAETDEKINRIQVTTSEISLVAGAPS---GDCCKNDANCDFSGDDGYAKD 1536
Db 1079 IQVQOSDNPE-----ETWNITVPGQSHVNVTLKANTPNVTLGVIRGYRTKPLY 1130
Qy 1537 NTPSSLAVCADGELYADL-----GNIRIRFKKK---PP-----L 1570
Db 1131 VETITGAHPVBGLTSDTIPESFNLSWTTNGDFDAFTIIDSRLLEPEMFNISGNS 1190
Qy 1571 NTQMYELSPIDQELYFTTGKLY-----TQSLPTGCDLYNFTYTGDDITL 1620
Db 1191 RTAHISGLSPSTDFIVLYGIS--HGFRQTCAISAAATTEAPEVDNLLVSDATPDGFRLT 1248
Qy 1621 ITDNG-----NNVNRDSTGMLNVLVP-----DQVYVWTMGT 1656
Db 1249 WTADDGVDFSVLKIRDKESD--PLELIVPGHERTHDITGLKEGTEYEIELYGVSSGR 1306
Qy 1657 NS-ALKSVTTQGHELAMMTYHGN-SGLLATKSNEN---GWTTFYEYDSFGRLTNVTFT 1710
Db 1307 RSQPINSVAT-----TVGSPKGLSFSDITENSARVSWTP-----PR 1343
Qy 1711 GOVSSFRSDTSSVHVQVETSSKDDVTITTLNLSAGAFYTLQDQVRNSYIIGADSLRL 1770
Db 1344 SRVDSYRVS-----YVPITGTEPNVTVDGSKTR---TKLVLP-----GVDYNVNI 1388
Qy 1771 LLANGHEVALQTEPHLLAGVNTVGNKN--VLPIDNGLNVLVWR-----QRK 1817
Db 1389 ISVKGF--ESEP--ISGLTKALDPSGLVNMVNTDSEALATWOPATAANDVNIYVS 1443
Qy 1818 EQARGQVT--VFGRRLRVHNRNLLSLDFDRVTEKIYDDHRKFTIRILYDQA--GRPSL 1873
Db 1444 SEDEPEVTQVMSGTVE-YDLNGLRPAATEYTLRVHAVKDAQKSETLSTQFTGLDAPKDL 1502
Qy 1874 WSPSSRLNGVNTYSPGYTAGTQRMSEMYDQAGRITSIFADGTWSTYILEKSM 1933
Db 1503 SATEVQSETAVITWRPP-----RAPVTDYLLAYESIDGRVKEVILDPETTSYLTLE--- 1553
Qy 1934 VLLHRSQRYIFPFQDKNRLSSVTMPNVARQTLETIRSVGYERNIYQPPEGNASVIQDFT 1993
Db 1554 ---LSPSTQTVVKLQALSR---SMRSKMIQTVFTTTGLLY-----PYPKDCSQALLNGEV 1602
Qy 1994 EDGHLHTFYLGTR 2008
Db 1603 TSG--LYTIYLNDR 1615

RESULT 12

T09070
probable tenascin X - mouse
C;Species: Mus musculus (house mouse)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jan-2000
C;Accession: T09070
R;Rowen, L.; Mahairas, G.; Qin, S.; Ahearn, M.E.; Dankers, C.; Lasky, C.; S.
submitted to the EMBL Data Library, October 1997
A;Description: Sequence of the mouse major histocompatibility locus class III region.
A;Reference number: Z16543
A;Accession: T09070
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-4006 <ROW>
A;Cross-references: EMBL:AF030001; NID:g2564945; PID:g2564958
C;Genetics:
A;Gene: TNX
A;Map position: 17

A;Introns: 124/1; 735/1; 773/3; 826/1; 914/1; 1037/1; 1135/1; 1232/1; 1329/1; 1440/1; 1
019/1; 3115/1; 3208/1; 3402/1; 3405/1; 3517/1; 3558/1; 3606/1; 3646/1; 3694/1; 3737/3;
C;Superfamily: tenascin-X; EGF homology; fibrinogen beta/gamma homology; fibronectin ty
C;Keywords: extracellular matrix
F;422-448/Domain: EGF homology <EGF>
F;826-906/Domain: fibronectin type III repeat homology <3FR>
F;3785-3997/Domain: fibrinogen beta/gamma homology <FBG>

Query Match 4.2%; Score 626; DB 2; Length 4006;
Best Local Similarity 22.4%; Pred. No. 1.2e-25;
Matches 359; Conservative 163; Mismatches 585; Indels 498; Gaps 77;

Qy 566 CPSNCYGVNGDCISGTCCHFLGFLGPDGRASCPVLCSNGQVMKGRCLCHSGWGKAECDV 625
Db 422 CPDRCRGRGREDGVCVCHAGYSBEDCVRSRCPGDCRGNCSGRVCVFGYTRDGC- 480
Qy 626 PTNOCIDVACSNHGTCITGTCINPGYKESCEEVDMDPTCSGRGVCVRGECHECFVGMG 685
Db 481 -TRAC-PGDCRGRGRCVGRVCNPGFTGEDCGSRRC-PGDCRGGHCHENGVCVCAVGYS 537
Qy 686 CTNCETPRATCLDQCSHGHTFLPDTGLCSDPSTWTHDCSIEICAADCGGHGVCVGGTCR 745
Db 538 GDCCST--RSCPSDCRGRGQCL--NGLCEDEGYSGEDCGIRRCPRDCSQHGVQDGLCM 593
Qy 746 CEDGWMGAACDQACHPRCABHGTCRCQKCECSQWNGEHCCTIAHYLDVVVKESGCPGLCN 805
Db 594 CHAGVAGEDCSIRTCPADCRRRGRCEGRCVCNPGYTGPA-ATRTCPADCR 644
Qy 806 NNGRCTLNLNWHVCVQLGWRGAGCDTSMETA-----CGDSKNDGDLVDCMD----P 855
Db 645 GRGRVCQGV---CMCYVYSGDEDCGQEPASACPGCG-PRELCRAGQCVCVGRFRGP 699
Qy 856 DCLQPL--CHINPLCL-----ODLRVAVT-----PTTLDLSWL 848
Db 700 DCAITCPGDCRSRGECLQRCVCOEGYAGDDCCEIPIQNMHMLLEETVTRWTRA 759
Qy 873 NPLDIIQETQVPSVQQNLHSHFYDRIKFLVGR-DSTHIIPGENPDGHCACVIRQVMTS 931
Db 760 PGVDVAYEQIPMTEGVSPPTARVPSSAYSAYDQRLAPGD-----YQVTVRALRGTS 814
Qy 932 DGTPLVGNISFVNNPLFGYISRDGSPDLVTNGGISILIRFEAPITQEHHTLWLPK- 990
Db 815 WQPPASKTITMIDGP-----ODLRVAVT-----PTTLDLSWL 848
Qy 991 -----DRFFVMTIIMRHEEN-EIPSCDLSNFARPNVWSPSLTSF----- 1031
Db 849 RQAEVDFVSVYSAGNRVLEVP-----PEADRTQLTDLMPGVEYVVTVA 897
Qy 1032 -----ASCAEKGFIVPEIQALQEEISISCKMRLSYLSSRTPGYKSVLRISLTHP 1082
Db 898 ERGHAVSYPASTRANTGSLPSGLLEATDEPPPSG-----PSTTQGAQAPILILEHH- 948
Qy 1083 TIFPNLMKVHLMVAVEGRLEFRKWFAPAAAPDLSYF-----IWDKTDV----- 1123
Db 949 --PLGELKV-LGRDKAGRLSVAN-TAQDPSFAHQRLRMOVAEGPWAHEILLPGDVOQALV 1004
Qy 1124 -----YNQKVGLSE-AFVSVGYEYESCPDILMEKRTTVLQGYEIDASKLGGWSLD 1174
Db 1005 PPPPGAPYKFLHGITPGKISVPIYQIGMD-----RAEQPGKPSQPRLGELTVP 1058
Qy 1175 KHHALNIOGILHKGNGENQFVS-----QQPPVIGSIMGNGRRSISCSFSCNGLADGN 1227
Db 1059 ---GLTSSLLHWTVPGEFDSFLIQYKDXDQOASVEGFPQRSTPLS-----GLEPGR 1110
Qy 1228 K-----LLAPVALTCGSD-----GSLYVVG- -FNYIRR 1253
Db 1111 KYKFLYGLIGKKRGPLMAEAKILIQSDPDGPPRGLGELWVDPTPHSLHLSWTVLGG 1170
Qy 1254 IPPSGNVTNILELRNKDFR-----HSHSPAHKY- -LA 1284
Db 1171 QFDS-----FVQYRDKGQPRVWVEGPDRSVWISPLDPNKRKFTFLFGIANKKRYGSLT 1226
Qy 1285 TDPMSGAVFLSDSNS-RRVFXIKSTVVVKDLVKNSVVA-GTGD-----QCL 1329

Db 1227 ADGTTAAETKESESPELCELVTGTPDLSRLSWTVARGPDSFVILYKDAQGPQSV 1286
Qy 1330 PFDDTRCGDGKATEATLTPRGTIVDKFGLIYFVDTMIRRDQNGIISTLLGSDNLT 1389
Db 1287 PIE-----GDNEVTVPGLESNRYKKNLYGL-----RGQRVPVSVV-----AKT 1328
Qy 1390 ARPLSCDSVMDISOVRLWETDPLAINPMDNSLVLD---NNVLQ--ISENH-----QV 1438
Db 1329 APPVDVDPSPSPSTEAP-EPPEKPLRELIVTDTTPNSLSLWTVSEGGQDFQVQY 1387
Qy 1439 RIVAGRPMHCQVQ-----IDH-FLSKVAIHATLESATLAVSHNGVLVIAT 1486
Db 1388 KNGDQPKVVRVPGHEDQVITISGLEPDKYKKNLYGIH-----SGQRVPVSVVGTAA 1441
Qy 1487 DEKKINRIQVTS-----GETSLVAGAPSGCDKNDANDCFSS-----GD-DG 1529
Db 1442 EEPSPSTESTAPETPEPLGELIVTGGSP-----DSLSLWTVPGQDFDS 1490
Qy 1530 YAKDAKLNTSPSLAVCADGELYVADLG-----NIRIRFIRKNKPLNTQMYELS-----S 1580
Db 1491 FTQYK-----SSDGRPQVMRVGQREIVVRDLDPGHKY--KKNLYGLHEHRYG 1539
Qy 1581 PIQELVFTTGGKLYTQSLPTGDIYLYNFTYTGDDGITLITDNGMVMVNRDSTGMPL 1640
Db 1540 PVS-----TVGTDTQK-PEEKPPLEPLBELVTG-----VTPNS-----VSL 1576
Qy 1641 WLVPDQ-----VYVMTGNTSALKSVTTQGHSLAM-----MTYHNSG----- 1680
Db 1577 SWIVPEGQDFSVVQYQDRDQPREVPLTAEQREVTVDLDLPARIYKMFYGLHGRRVG 1636
Qy 1681 -----LLATKNE-----NGWTPYFY--DSFGLNTVPTTQGVSSFR---SD 1719
Db 1637 PLSVSAMTAPAPATEASKLPEPRLGELVTDVTPDSVGLW--TVPEGEFDSFTVQYXD 1694
Qy 1720 TDSVHVQVETSSKDDVTITNLSAGAFVTL---QDVRSNYIGADGSLRLLLANM 1776
Db 1695 RDGQPHVVSVAADLEVTI-PLGLEPKYKFLFGVQDGKRSOV-----SV 1740
Qy 1777 EVALQTEPHLAGTVNPTGKRVNVTIPDNLNLVWR-----QRKEARQGV 1825
Db 1741 EAKTATHTDTPSG-APPRLGELWVTDPTDLSRL-SWTVPEGHFDSFVQVQKXGPRVV 1798
Qy 1826 VFGRRVHRNLLSDFDVTETKTYDDHRTFLRLLYDQGR 1870
Db 1799 PVKQER-----TATITALDGRKYRF-LYGLLGR 1828

RESULT 13

A40701
C;Species: human
C;Species: Homo sapiens (man)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Dec-1999
C;Accession: A40701; A33725; C42175
R;Bristow, J.; Tes, M.K.; Gitelman, S.E.; Mellon, S.H.; Miller, W.L.
J. Cell Biol. 122, 265-278, 1993
A;Title: Tenascin-X: a novel extracellular matrix protein encoded by the human XB gene
A;Reference number: A40701; UID:93300909; PMID:7686164
A;Accession: A40701
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-3566
A;Cross-references: EMBL: X71937
R;Morel, Y.; Bristow, J.; Gitelman, S.E.; Miller, W.L.
Proc. Natl. Acad. Sci. U.S.A. 86, 6582-6586, 1989
A;Title: Transcript encoded on the opposite strand of the human steroid 21-hydroxylase/c
A;Reference number: A33725; UID:89367293; PMID:2475872
A;Accession: A33725
A;Molecule type: mRNA
A;Residues: 2748-3199, 'V', 3201-3298, 'E', 3299-3314, 'G', 3316-3566
A;Cross-references: GB:M25813; NID:g183069; PIDN:AAA35884.1; PID:g183070
R;Matsumoto, K.; Arai, M.; Ishihara, N.; Ando, A.; Inoko, H.; Ikemura, T.
Genomics 12, 485-491, 1992

A;Title: Cluster of fibronectin type III repeats found in the human major histocompatibi
enasin.
A;Reference number: A42175; UID:92217969; PMID:1373119
A;Accession: C42175
A;Molecule type: DNA
A;Residues: 1849-1936
A;Experimental source: clone 3.9kF3-1
A;Note: sequence extracted from NCBI backbone (NCBIP:95694)
C;Genetics:
A;Gene: GDB:TXNA; D6S103B; TNX; XA; XB
A;Cross-references: GDB:568487; OMIM:600261
A;Map position: 6p21.3-6p21.3
C;Superfamily: tenascin-X; EGF homology; fibronogen beta/gamma homology; fibronectin typ
C;Keywords: extracellular matrix; glycoprotein
F;435-461/Domain: EGF homology <EGF>
F;748-828/Domain: fibronectin type III repeat homology <3F1>
F;829-856/Domain: fibronectin type III repeat homology #status atypical <3F2>
F;873-933/Domain: fibronectin type III repeat homology <3F3>
F;975-1055/Domain: fibronectin type III repeat homology <3F4>
F;1078-1158/Domain: fibronectin type III repeat homology <3F5>
F;1167-1247/Domain: fibronectin type III repeat homology <3F6>
F;1248-1317/Domain: fibronectin type III repeat homology #status atypical <3F7>
F;1323-1403/Domain: fibronectin type III repeat homology <3F8>
F;1412-1492/Domain: fibronectin type III repeat homology <3F9>
F;1510-1590/Domain: fibronectin type III repeat homology <3F10>
F;1618-1676/Domain: fibronectin type III repeat homology #status atypical <3F11>
F;1678-1749/Domain: fibronectin type III repeat homology #status atypical <3F12>
F;1751-1831/Domain: fibronectin type III repeat homology <3F13>
F;1849-1929/Domain: fibronectin type III repeat homology <3F14>
F;1955-2035/Domain: fibronectin type III repeat homology <3F15>
F;2081-2141/Domain: fibronectin type III repeat homology <3F16>
F;2167-2246/Domain: fibronectin type III repeat homology <3F17>
F;2274-2354/Domain: fibronectin type III repeat homology <3F18>
F;2382-2462/Domain: fibronectin type III repeat homology <3F19>
F;2488-2568/Domain: fibronectin type III repeat homology <3F20>
F;2584-2664/Domain: fibronectin type III repeat homology <3F21>
F;2677-2757/Domain: fibronectin type III repeat homology <3F22>
F;2771-2851/Domain: fibronectin type III repeat homology <3F23>
F;2878-2958/Domain: fibronectin type III repeat homology <3F24>
F;2977-3067/Domain: fibronectin type III repeat homology #status atypical <3F25>
F;3078-3159/Domain: fibronectin type III repeat homology <3F26>
F;3157-3247/Domain: fibronectin type III repeat homology <3F27>
F;3255-3334/Domain: fibronectin type III repeat homology <3F28>
F;3349-3557/Domain: fibronogen beta/gamma homology <FBG>

Query Match 4.0%; Score 592; DB 1; Length 3566;
Best Local Similarity 21.4%; Pred. No. 7.8e-24;
Matches 364; Conservative 182; Mismatches 549; Indels 606; Gaps 86;

Qy 562 SVDNCPNCGNGPCISGTCHCFLGFLGPDGGRASCPLVCSGNGQYMKGRCLCHSGMKA 621
Db 369 STRCPDRGRGRCEDEGEICITGYSGDDGCVRSRCPGDCNQGRCEDEGRVCWPGTGT 428
Qy 622 ECDVPTNQCIDVACSNHGTCTITGTCINPGYKSGESCSEVDMPTCSGRGVVGRGCHCF 681
Db 429 DCG---SRACPRDCRGRGRCENGVCVNCAGYSGEDCGVRSC-PQDCRGRGRCESGRCMW 484
Qy 682 VGMGWTNCET-----PRATCLD-----QCSGHGTFLPDTGL 712
Db 485 PGYTGRDCGTACPGDCRGRGRCVDRGVNCPNPTGDCGSRRCRPGDCRCHG--LCEGV 542
Qy 713 CSCDPSWTGHDSCSTETCAADCGGHGVCVGGTCTCRCEGMMGAACDQRAHPRCAHGTCD 772
Db 543 CVCDAGYSGEDCSFRSCPGGRCRGQCLDGRVCEDGYSGEDCGVRQCPNDCSQHGVQCD 602
Qy 773 GKCCSCPGWNGEHCETIAHYLDVVKCEPGLCNNGRCITLDNGHVCVQLGWRGAGCDT 832
Db 603 GVCICWEGYVSEDCSI-----RTPCNSNCHRGRC-----BEGRLCLDPFGTGTGTCAT 649
Qy 833 SMETACGDSKNDGDGLVDQMDPCCLQP--LCHIN-----PLCLGSRNPLD 877
Db 650 RMCPA-----DCRGRGRCVQVCLCHGVGGEDCGEBPPASACFGCGPRE 696

Qy 878 LIQETVPVSVQNLHSGFYDRIKELVGRD-STHIIIPGE-----NPPDGGHACVIRGOVMTS 931
 Db 697 LCRAGQVC-----VEGRPDCAIOTCPGDCRGRGECHDG--SCVCKDYGAGE 743
 Qy 932 DGTPLVGNWISFVNNPLFGYISRQDGSFLVINGGISILRPERAPFITOEHTLWLPW- 990
 Db 744 DCEVI-----DGPQDL-----RVSVTPT-TLELGLW 770
 Qy 991 -----DRFVM-----ET-----IIMRHEENIPSCDLS 1014
 Db 771 RPAQEVDFVSVYSADNQVRULEVPPETDGLTLLDMPGVVYVVTVAERGRAVYPAS 830
 Qy 1015 NFARNPFWAS-PGPLTSSFASSCAEKGPVPEIQALQEEISISCKKRLS--YLSSRTPGY 1071
 Db 831 VRANTDGRKYKFLVYGFVK-KHGGFLVAEAKILPQSDPSPTPRLGNLWTDPTP-- 887
 Qy 1072 KSVLRISLTHTPTTFNLMKHL-----MVAVEGRLEFRKWFAPAADLSYFIWDKTDV 1123
 Db 888 -DSLHLWSWTVPGEQDFDMVQYDRDRGPQVVEGP-ERSFVVSLLDPDKYRFTLFGI 945
 Qy 1124 YNOKVFLSBAFVSVGYEYES-----CPD--LILWEKRTTVLOG 1160
 Db 946 ANKKRYOPLTADGTTAPERKEEPRPEFLQPLLGLTGTGTPDSLSLSW-----TVAQG 1001
 Qy 1161 -----YEIDAS-----KLGWSLDKHALNIQSGILHKGNGENQFVS 1197
 Db 1002 PFDSEFMVQYK-DAQQOPQAVPVAGDENEVTPGLDPRKYMNLGLGRQRVGPESVA 1060
 Qy 1198 QBPVIGSINGERRISCPSCNGLADGNKLAPVALTCSPGSLVYGVDFNIRIF-- 1255
 Db 1061 KTDG-----QPHRTTEAPE-----SPEPLILGELTVTSGSPDSL--SLFWT 1100
 Qy 1256 -PSGNVTNI-LELRNCKDFR-----HSHSPAKEYL----- 1283
 Db 1101 VPQSPDFTVQYKDRDRPRAVRVGKSEVTVGGLPEGHKYMHLGLHEGORVGPVS 1160
 Qy 1284 -----ATDPMGAVFLSDNSRRV-----FKIKSTTV-VKDLVKNSEVVAAGTGOCL 1329
 Db 1161 AVGLTALPRGLTVDITPDSVGLSWTVPEGEFDSFVQYKDRDQGPQVPEVAADQ-- 1218
 Qy 1330 PFDTRCGDGGKATEALTNPRTGTVTKFGLIVFDG----- 1366
 Db 1219 -----REVTPLESTKYKFLFLGIDGKRSPSVSEAKTAILSWTVPEAS 1265
 Qy 1367 ---TMIRRIDONG-----IISTLLGNDLTSARPUSCD- 1396
 Db 1266 STQKVQFKDKDGPQVVPVEGHERSVTVPLDAGRKYRFLLYGLLKK-----RHGPLTADG 1322
 Qy 1397 -----SYMDISQ--VRLEWPTDLAINPMDNSLYVLNNVVLQISENHQVRI--VAGR 1444
 Db 1323 TSGEGHSLQVTTVTQNSVGLSWTVPEG-----QFDSFVQYKDRDQGPQVPEGS 1373
 Qy 1445 PMHQVPGID---HFLLSKVAIH--ATLESATALAVS-----HNGVLYIAETDEKKINRIR 1495
 Db 1374 LREVSFGLDPAHYRKLILYGLHKGKRVGPISAVAITAPEPHLGELTVBEATSHTLH-LS 1432
 Qy 1496 QVTTSGISLVAGAPSGCDKNDANCDPFGDDGYAKDAXLNTPSSLAVCADGELYADL 1555
 Db 1433 WMVTEGEFD-----SFEIQYTD-----RDGQLQWRI 1459
 Qy 1556 G---NIRIRIRKKNKFLNTQNMVYELSPIDQELVLDFTTGKH---LYTQSL---PTGD 1605
 Db 1460 GGDNDITLSGSDHRYLVT-----LYGF-SDGKHGVPHVHEALTEPTA- 1504
 Qy 1606 YLYNFTYTD-----GDITLITDNNGMVNVARDSTGMPLMLVVPDQV-VWVTMGTN 1657
 Db 1505 -----TPEPIKPRLGELT-VTDATPDSLSL-----W-TVPEGFDFHFLVOYRN 1547
 Qy 1658 --SALKSVTTQGEH-----LAMTYHGS-----GLLATKSNENGWT 1692
 Db 1548 GDGQKAVRGPGBEGEVTISGLEPDHKYKMNLYFGHGQRMGPVSVGVVTAASEETPSPT 1607
 Qy 1693 -----TFVEYDSFGRLT-----NVTPTGQVSFR-----SDTDSVHV--Q 1727

Db 1608 EPSMEAPFAPEEPLLGELTVTGSSPDLSLSWTVPQGRFDSFTVQYKDRGRPQVRVG 1667
 Qy 1728 VETSSKD-----DVTITNLSAS-----GAF-YTLLQDVRNSYIIGADGS 1767
 Db 1668 PEESPDAPLAKRLGQMTVRDIT-SDLSLSWTVPEGFDFHFLVQ-----FKNGDQ 1719
 Qy 1768 LRLILLANMEVALQTEPHLLAGTVNPTVGKRNVTLPIDNGLNLVEMRQRKEQARGQVTF 1827
 Db 1720 PKALRVPHEDGASEE-----TPSPTEPSMEAEPEP-----PEEPLLGELIVT 1761
 Qy 1828 GRRLRVHNRNLLSDFDRTVTEKIYDHRKFTLRLYDQAGRPSLMSPSRLNGVNTY 1887
 Db 1762 G-----SSPDSLSLSW--TVPQGRFD--SFTVQ-YKDRDGRPQV--VRVGGESEV 1805
 Qy 1888 SPGG-----YIAGIQRG 1899
 Db 1806 TVGGLEPGRKYMHLHYGLHEG 1826

RESULT 14
 A:3902
 tenascin - eastern newt (fragment)
 C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 12-Feb-1999
 C:Accession: A43902
 R:Onda, H.; Poulin, M.L.; Tassava, R.A.; Chiu, I.M.
 Dev. Biol. 148, 219-232, 1991
 A:Title: Characterization of a new tenascin cDNA and localization of tenascin mRNA dur
 A:Reference number: A43902; MUID:92038434; PMID:1718799
 A:Accession: A43902
 A:Molecule type: mRNA
 A:Residues: 1-647 <OND>
 A:Cross-references: GB:M76615
 A:Note: sequence extracted from NCBI backbone (NCBIN:64543, NCBIP:64547)
 C:Superfamily: tenascin; EGF homology; fibronogen beta/gamma homology; fibronectin type
 C:Keywords: extracellular matrix; glycoprotein; tandem repeat
 F:287-313/Domain: EGF homology <EGF>
 F:346-427/Domain: fibronectin type III repeat homology <FNI>
 F:435-517/Domain: fibronectin type III repeat homology <FN2>

Query Match 3.9%; Score 577.5; DB 2; Length 647;
 Best Local Similarity 27.8%; Pred No. 3.2e-24;
 Matches 161; Conservative 70; Mismatches 208; Indels 141; Gaps 25;

Qy 566 CPSNCGYNGDCISGTCFCHGFLGPDGCRASCPVLCSNGQYMKGRCLCHSGWKGAECDC- 624
 Db 70 CPNDCFDGRGRCINGVCFCEGFTGEDCGELTCCPNNCNRGRVCNGLCVCDGFGDGCSE 129
 Qy 625 --VPTNQCIDVACSNHGTCTTCICNPGYKGESCEEVDMDPTCSGRGVCVRGECFCFV 682
 Db 130 LRCFND-----CNDGRGRCVCKVCEGFMGEDCADLRCPN-DCNNRGRVNGQCVDE 182
 Qy 683 GWGGTNCETPRATCLDQCSGHGTFLPDTGLCSCDPSWTHDCSIBICAADCGHGVGVG 742
 Db 183 GFMGEDCSDLR--CPGDCNRRGCV--NGQCVCDEGFRGEDCGELRCPDDCNRRGVCVNG 238
 Qy 743 TCRDEGWMGAACDORACHPRCAEHGTCDGCKECSGPNWGEHCTIAHYLVRVKEGCGP 802
 Db 239 QCICDEGFMGNCGLRCPNDCNRRGVCNNGQICDDGFKGEDCS-----ELRCPD 289
 Qy 803 LCNNGRCFTLDLNGHVCQLGWRGAGDTSMETACGDSKNDGDLVDCMDPDCCLOPL 862
 Db 290 DCDNRGRC---ING-QVCVCAEGFTGENDCS---LAC-----LNNCDRGLCVNGQ 332
 Qy 863 CHINPLCLG-----SPNPLDI-----IQE---TQVPVSOQNLHSF 894
 Db 333 CVCEEGFLGEDCSEVSP-PKDLTVTDVTQSVNLEWANEKVTLEYLTIPTSPGL--- 388
 Qy 895 YDRIKFLVGRDST-----HIIPGENPFDGGHACVIRGOVMTSDGTPLGVNISFVNNPLF 949
 Db 389 --ELDFRVPGDQTTATIQELEBPGVYFVRVFA-ILRNQ-----RSIPV----- 428

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 16:15:08 ; Search time 85 seconds
(without alignments)
10278.469 Million cell updates/sec

Title: US-10-029-020-14

Perfect score: 14887

Sequence: 1 MDVKRKYPSLRRDAER.....ELSDGANNHFVQSEMGR 2769

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 segs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: SP archaea.*
- 2: SP bacteria.*
- 3: SP fungi.*
- 4: SP human.*
- 5: SP invertebrate.*
- 6: SP mammal.*
- 7: SP mhc.*
- 8: SP organelle.*
- 9: SP phage.*
- 10: SP plant.*
- 11: SP rodent.*
- 12: SP virus.*
- 13: SP vertebrate.*
- 14: SP unclassified.*
- 15: SP_rvirus.*
- 16: SP_bacteriap.*
- 17: SP_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14529	97.6	2771	11 Q9WTS7	Q9WTS7 mus musculus
2	14269	95.8	2825	11 Q70465	Q70465 mus musculus
3	12033.5	90.8	2824	13 Q9W7R3	Q9W7R3 brachydanio
4	10227	68.7	2715	11 Q9WTS6	Q9WTS6 mus musculus
5	9968	67.0	2192	13 Q804R1	Q804R1 brachydanio
6	9642	64.8	2705	13 Q9W6V6	Q9W6V6 gallus gall
7	9616.5	64.6	2802	13 Q9DER5	Q9DER5 gallus gall
8	9610.5	64.6	2764	11 Q9WTS5	Q9WTS5 mus musculus
9	9607	64.5	2731	11 Q9WTS4	Q9WTS4 mus musculus
10	9601	64.5	2765	11 Q9R1K2	Q9R1K2 rattus norv
11	9501	63.8	2725	4 Q9UKZ4	Q9UKZ4 homo sapien
12	9303	62.5	2590	13 Q9W7R4	Q9W7R4 brachydanio
13	9145	61.4	2346	11 Q9JLC1	Q9JLC1 mus musculus
14	8059.5	54.1	2144	4 Q9ULU2	Q9ULU2 homo sapien
15	7063	47.4	1828	11 Q80TD2	Q80TD2 mus musculus
16	6774	45.5	1769	4 Q9P273	Q9P273 homo sapien

17	6189	41.6	1198	11 Q80TF5	Q80TF5 mus musculus
18	5638	37.9	1086	4 Q9P2P4	Q9P2P4 homo sapien
19	5614	37.7	1071	4 Q723C7	Q723C7 homo sapien
20	3973.5	26.7	1045	4 Q9NVW1	Q9NVW1 homo sapien
21	3854	25.9	730	4 Q96MS6	Q96MS6 homo sapien
22	3822	25.7	2731	5 Q61307	Q61307 drosophila
23	3820.5	25.7	2731	5 Q9VNU6	Q9VNU6 drosophila
24	3813	25.6	2731	5 Q18366	Q18366 drosophila
25	3804.5	25.6	2515	5 Q24551	Q24551 drosophila
26	3753	25.2	991	11 Q8CAT1	Q8CAT1 mus musculus
27	3720	25.1	964	4 Q9NV77	Q9NV77 homo sapien
28	3517	23.6	3004	5 Q24550	Q24550 drosophila
29	3515	23.6	3004	5 Q9VYN8	Q9VYN8 drosophila
30	3398.5	22.8	930	11 Q9JLC0	Q9JLC0 mus musculus
31	3111.5	20.9	849	4 Q9NT68	Q9NT68 homo sapien
32	2378	16.0	831	13 Q9PU49	Q9PU49 gallus gall
33	2351	15.8	625	4 Q96SV2	Q96SV2 homo sapien
34	2108.5	14.2	2560	5 Q21980	Q21980 caenorhabdi
35	2079.5	14.0	2531	5 Q8MP22	Q8MP22 caenorhabdi
36	1464	9.8	337	11 Q9R1K0	Q9R1K0 rattus norv
37	1336	9.0	278	13 Q9DEQ8	Q9DEQ8 gallus gall
38	1244	8.4	442	4 Q9NZJ2	Q9NZJ2 homo sapien
39	1129	7.6	536	11 Q8C9D2	Q8C9D2 mus musculus
40	1120.5	7.5	272	11 Q9R1J9	Q9R1J9 rattus norv
41	1093.5	7.3	266	11 Q9R1K1	Q9R1K1 rattus norv
42	1089	7.0	229	11 Q9QYZ1	Q9QYZ1 mus musculus
43	1036	6.9	1009	5 Q8MQJ6	Q8MQJ6 drosophila
44	1034.5	6.9	329	11 Q8BSL5	Q8BSL5 mus musculus
45	943	6.3	184	4 Q9Y4S2	Q9Y4S2 homo sapien

ALIGNMENTS

RESULT 1
QWTS7
ID Q9WTS7 PRELIMINARY; PRT; 2771 AA.
AC Q9WTS7
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Ten-M4
GN NRGI OR OD24 OR TEN-M4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c; TISSUE=Brain;
RA Ohashi T., Zhou X., Feng K., Richter B., Moergelin M., Perez M.T.,
RA Su W., Chiquet-Ehrismann R., Rauch U., Faessler R.;
RT "Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
RT proteins expressed in many tissues."
RL J. Cell Biol. 0:0-0(1999).
DR EMBL; AB025413; BAA77399.1; -.
DR HSP; P16109; 1P5B.
DR MGD; MGI:96083; Nr1.
DR MGD; MGI:2447063; Odz4.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR006530; YD.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF01436; NHL; 6.
DR Pfam; PF05593; RNS_repeat; 5.
DR SMART; SM00181; EGF; 6.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 7.
DR PROSITE; PS00022; EGF_1; 8.
DR PROSITE; PS01186; EGF_2; 7.
KW EGF-like domain.
SQ SEQUENCE 2771 AA; 308497 MW; 1492E1E1A0DBF0C CRC64;

Query Match				97.68; Score 14529; DB 11; Length 2771;			
Best Local Similarity				97.08; Pred. No. 0;			
Matches 2688; Conservative				40; Mismatches 41; Indels 2; Gaps 1;			
Qy	1	MDVKRPYSLRRDAERYTSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60				
Db	1	MDVKRPYSLRRDAERYTSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60				
Qy	61	KDVPQAEBCFCANFTLREILGLEVTPPHGTLYRTDILGUPQCGSYNGASDADMEAD	120				
Db	61	KDVPQAEBCFCANFTLREILGLEVTPPHGTLYRTDILGUPQCGSYNGASDADMEAD	120				
Qy	121	TVLSPEHPVLWGRSTRSGSSCLSRANSNLTLTDEHENTETDHPGLQNHARLRTPP	180				
Db	121	TVLSPEHPVLWGRSTRSGSSCLSRANSNLTLTDEHENTETDHPGLQNHARLRTPP	180				
Qy	181	PPLSHAHTNQHAASINSUNRGNFTPRNSPAPTDLHSLGEPAGGAQBAHAQBNWL	240				
Db	181	PPLPHAHTNQHAASINSUNRGNFTPRNSPAPTDLHSLGEPAGGAQBAHAQBNWL	240				
Qy	241	LNSNIPLERNLKGQPLGTQDNLEMDTLGASRHDGAYSDGHFLFKPGGTSPLFCCTS	300				
Db	241	LNSNIPLERNLKGQPLGTQDNLEMDTLGASRHDGAYSDGHFLFKPGGTSPLFCCTS	300				
Qy	301	PGYPLTSSTVYSPPPRLPRSTFARPAFNLKKPSKYNMKCAALSIAIVISATLVILLAYF	360				
Db	301	PGYPLTSSTVYSPPPRLPRSTFARPAFNLKKPSKYNMKCAALSIAIVISATLVILLAYF	360				
Qy	361	VAMHFLGNLHLOPMEG-OMYEITETASSMVPDVSLSGCGTGLTPDRKKGITTE	418				
Db	361	VAMHFLGNLHLOPMEGOMYEITETASSMVPDVSLSGCGTGLTPDRKKGAAE	420				
Qy	419	GKPSFFPEDSFIDSGEIDVGRASQKIPPGTFWRSQVFIIDHPVHLKFVSLGKAALVGI	478				
Db	421	GKPSFFPEDSFIDSGEIDVGRASQKIPPGTFWRSQVFIIDHPVHLKFVSLGKAALVGI	480				
Qy	479	YGRKGLPPSHQTQDFVELLDGRLLTQEARSLBGTFRQSRGTVPSSSHETGFIQYLDGSI	538				
Db	481	YGRKGLPPSHQTQDFVELLDGRLLTQEARSLBGTFRQSRGTVPSSSHETGFIQYLDGSI	540				
Qy	539	WHLAFYNDGSESVSFLTAIESVDCNSCNGDCISGTCCHCFGLGFLGDCGRASCP	598				
Db	541	WHLAFYNDGSESVSFLTAIESVDCNSCNGDCISGTCCHCFGLGFLGDCGRASCP	600				
Qy	599	VLCSGNGQYMKRCLCHSGWKGAECVPTNQCIDVACSNHGTCTGTCTCNPGYKGESCE	658				
Db	601	VLCSGNGQYMKRCLCHSGWKGAECVPTNQCIDVACSNHGTCTGTCTCNPGYKGESCE	660				
Qy	659	EVDCMPTCSGRGVCYRGECHCFVWGNGTNCETPRATCLDQCSGHGTFLPDTGLCSCDPS	718				
Db	661	EVDCMPTCSGRGVCYRGECHCFVWGNGTNCETPRATCLDQCSGHGTFLPDTGLCNCDDPS	720				
Qy	719	WTGHDCSIETCAADCGHGVGVGTCRCEDGWMGAACDQACHPRCAEHGTCTRDGKCECS	778				
Db	721	WTGHDCSIETCAADCGHGVGVGTCRCEDGWMGAACDQACHPRCAEHGTCTRDGKCECS	780				
Qy	779	PGNNGEHCCTIAHLDVWVKEGCEGLCNGRCCTLDLNGHVCOLGWRGAGGCTSMETAC	838				
Db	781	PGNNGEHCCTIAHLDVWVKEGCEGLCNGRCCTLDLNGHVCOLGWRGAGGCTSMETGC	840				
Qy	839	GDSDKNDGDGLVDCMPDCCQLPCLHINPLCLGSPNPLDIIIOETQVPVSQQNLHGFYDRI	898				
Db	841	GDSDKNDGDGLVDCMPDCCQLPCLHINPLCLGSPNPLDIIIOETQAPVSQQNLNFPYDRI	900				
Qy	899	KFLVGRDSTHIIIPENPFPGGHACVIRGQVMTSDGTPLVGVNISFVNNPLFGYTTSRQDG	958				
Db	901	KFLVGRDSTHIIIPENPFPGGHACVIRGQVMTSDGTPLVGVNISFVNNPLFGYTTSRQDG	960				
Qy	959	SFDLVNTGGSIILTRPERAPFTQEHHTLWLPDRFPVMEITIIMRHEENIEIPSCDLSNFAR	1018				
Db	961	SFDLVNTGGSIILTRPERAPFTQEHHTLWLPDRFPVMEITIIMRHEENIEIPSCDLSNFAR	1020				
Qy	1019	PNPVSFSPITSPASSCAEKGPIVPEIQALQBEIIVAGCKMRLSYLSRTPGYSKVLRLIS	1078				

Db	1021	PNPVSFSPITSPASSCAEKGPIVPEIQALQBEIIVAGCKMRLSYLSRTPGYSKVLRLIS	1080				
Qy	1079	LTHFTIPFNLMKVHLMVAVEGRFRKWFAAAPDLSYYFIWDKTDVYNQKVFGESEAFVSU	1138				
Db	1081	LTHFTIPFNLMKVHLMVAVEGRFRKWFAAAPDLSYYFIWDKTDVYNQKVFGESEAFVSU	1140				
Qy	1139	GYESCESPDILWEKRTTVLQGYEIDASKLGGWSLDKHALNIOGTLHXKNGENOPVSQ	1198				
Db	1141	GYESCESPDILWEKRTTVLQGYEIDASKLGGWSLDKHALNIOGTLHXKNGENOPVSQ	1200				
Qy	1199	QPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDYFNIRIRPPSG	1258				
Db	1201	QPPVIGSIMGNRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDYFNIRIRPPSG	1260				
Qy	1259	NVTNILELRNKDFRSHSPAHKYIYLATDPMGSAFLSDSNRRVFKIKSTVAVKDLVKNIS	1318				
Db	1261	NVTNILELRNKDFRSHSPAHKYIYLATDPMGSAFLSDSNRRVFKIKSTVAVKDLVKNIS	1320				
Qy	1319	EWAGTGDQCLPFDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRIDQNGII	1378				
Db	1321	EWAGTGDQCLPFDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRIDQNGII	1380				
Qy	1379	STLIGSNDLTSARPLSCDSVMDISOVRLEWPTDLAINPMDNSLYVDNNVLOISENHQV	1438				
Db	1381	STLIGSNDLTSARPLSCDSVMDISOVRLEWPTDLAINPMDNSLYVDNNVLOISENHQV	1440				
Qy	1439	RIVAGRPMHQVPGIDHFLLSKVAIHATLESATALAVSHNGVLVIAETDEKKINRIRQVT	1498				
Db	1441	RIVAGRPMHQVPGIDHFLLSKVAIHATLESATALAVSHNGVLVIAETDEKKINRIRQVT	1500				
Qy	1499	TSGEISILVAGAPSCDCKDANCDGSDGDAKADAKLNTPPSSLAVCADGELYVADLNI	1558				
Db	1501	TSGEISILVAGAPSCDCKDANCDGSDGDAKADAKLNTPPSSLAVCADGELYVADLNI	1560				
Qy	1559	RIRIRKKNKFLNTQNMVELSSPIDQELYLFDITGKHLYTQSIPGTDYLYNFYTGDDGI	1618				
Db	1561	RIRIRKKNKFLNTQNMVELSSPIDQELYLFDITGKHLYTQSIPGTDYLYNFYTGDDGI	1620				
Qy	1619	TLITDNNGNMNVNRDSTGMPLMLVDPDGVVYVMTGNSALKSVTTQGHELAMTYHGN	1678				
Db	1621	TLITDNNGNMNVNRDSTGMPLMLVDPDGVVYVMTGNSALKSVTTQGHELAMTYHGN	1680				
Qy	1679	SGLLATKSNENGWTTTFEYDSFGRLNVTPTQVSSFRSDTSSVHVQVETSSKDDVTI	1738				
Db	1681	SGLLATKSNENGWTTTFEYDSFGRLNVTPTQVSSFRSDTSSVHVQVETSSKDDVTI	1740				
Qy	1739	TTNLSASGAFYTLLODOVRNSYVIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKR	1798				
Db	1741	TTNLSASGAFYTLLODOVRNSYVIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKR	1800				
Qy	1799	NVTLPIDNGLNLEWROKQEQVTVFGRRURVNRNLLSDDFRVTREKTYDHRK	1858				
Db	1801	NVTLPIDNGLNLEWROKQEQVTVFGRRURVNRNLLSDDFRVTREKTYDHRK	1860				
Qy	1859	FTLRLVDQAGRLSLWSPSSRLNGVNTYSPGGVIAGIQGIMSERMEYDQAGRITSRIF	1918				
Db	1861	FTLRLVDQAGRLSLWSPSSRLNGVNTYSPGGVIAGIQGIMSERMEYDQAGRITSRIF	1920				
Qy	1919	ADGKTSYTYLEKSMVLLLSHQYQYTFEPDKNDRLSVTMPNVARQTLTIRSVGYRNI	1978				
Db	1921	ADGKTSYTYLEKSMVLLLSHQYQYTFEPDKNDRLSVTMPNVARQTLTIRSVGYRNI	1980				
Qy	1979	YQPEGNASVIQDTEDEHLLHTFYLGTRGVLYKYKLSKLAETLYDTTKSYTDETA	2038				
Db	1981	YQPEGNASVIQDTEDEHLLHTFYLGTRGVLYKYKLSKLAETLYDTTKSYTDETA	2040				
Qy	2039	GMLKATINLQNEGTCTTIRYQIGPLIDRQIFRTEEGVMNARFDYNDNSFRVTSQAVI	2098				
Db	2041	GMLKATINLQNEGTCTTIRYQIGPLIDRQIFRTEEGVMNARFDYNDNSFRVTSQAVI	2100				
Qy	2099	NEPPLPDLRYDDVSGKTEQPKFGVYYDINQIITTAWHTTKHFDAYGRKVEQYEI	2158				

Db 2101 NETPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQIITTAAMTHTKHFDAYGMKEVQVEI 2160
 QY 2159 FRSLMYMTVOYDNMGVVKELKVGPIYANTTRYSEYDADGLOQTVSINDKPLWRYSD 2218
 Db 2161 FRSLMYMTVOYDNMGVVKELKVGPIYANTTRYSEYDADGLOQTVSINDKPLWRYSD 2220
 QY 2219 LQNLHLSPGNSARLTPYRDIYRDIIRLGDVQYQMDDEGFLQRGGDIFEYNAGALLI 2278
 Db 2221 LQNLHLSPGNSARLTPYRDIYRDIIRLGDVQYQMDDEGFLQRGGDIFEYNAGALLI 2280
 QY 2279 KYNRAGSVRYRYDGLGRVSSKSHSHLQFFVADLTNPYKTVHLYNHSSSEITSY 2338
 Db 2281 KYNRAGSVRYRYDGLGRVSSKSHSHLQFFVADLTNPYKTVHLYNHSSSEITSY 2340
 QY 2339 YDLQGLHFAELSSGDEFYIACDNIPTPLAVFSGTGLMIKQILYAYGBIYDTPNFOI 2398
 Db 2341 YDLQGLHFAELSSGDEFYIACDNIPTPLAVFSGTGLMIKQILYAYGBIYDTPNFOI 2400
 QY 2399 IIGYHGGYDPLTKLVHMGSRDYDLVLAGRWTSYDHELMKHLSSNVMYKNNFI 2458
 Db 2401 IIGYHGGYDPLTKLVHMGSRDYDLVLAGRWTSYDHELMKHLSSNVMYKNNFI 2460
 QY 2459 SNSQDIKCFMTDVSNTLLTFGQLHNVIPGYPKPDMAMEPSYELIHTQMTQWDSKS 2518
 Db 2461 SNSQDIKCFMTDVSNTLLTFGQLHNVIPGYPKPDMAMEPSYELIHTQMTQWDSKS 2520
 QY 2519 ILGVOCEVQOLKAFVTLERFDOLYGTITSCQAPKTKKPASSGVSFGKGVKFKADGR 2578
 Db 2521 ILGVOCEVQOLKAFVTLERFDOLYGTITSCQAPKTKKPASSGVSFGKGVKFKADGR 2580
 QY 2579 VTDTIISVANEDGRVAAILNHAHYLENLHFTTIDGVDTHYFVKPSPSEGLAILGLSGGR 2638
 Db 2581 VTDTIISVANEDGRVAAILNHAHYLENLHFTTIDGVDTHYFVKPSPSEGLAILGLSGGR 2640
 QY 2639 RTLENGVNVTSQINTVNLGRTYTDIOYQALCALCNTRYGTTLDDEKARVLELARQA 2698
 Db 2641 RTLENGVNVTSQINTVNLGRTYTDIOYQALCALCNTRYGTTLDDEKARVLELARQA 2700
 QY 2699 VROAWREOQLREGEGLPAWTEGKQVLSGRVQGVGRFVIYVEQYBELSDSANNI 2758
 Db 2701 VROAWREOQLREGEGLPAWTEGKQVLSGRVQGVGRFVIYVEQYBELSDSANNI 2760
 QY 2759 HFMRQSEMGR 2769
 Db 2761 HFMRQSEMGR 2771

RESULT 2
 O70465 PRELIMINARY; PRT; 2825 AA.
 AC O70465; ID
 DT 01-AUG-1998 (TREMBlrel. 07, Created)
 DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE DOC4.
 GN NR01 OR OD24 OR DOC4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NIH/Swiss;
 RX MEDLINE=98315054; PubMed=9649432;
 RA Wang X.Z., Kuroda M., Sok J., Batchvarova N., Kimmel R., Chung P.,
 RA Zinszner H., Ron D.;
 RT "Identification of novel stress-induced genes downstream of chop."
 RL EMBO J. 17:3619-3630(1998).
 DR EMBL; AF059485; AAC31807.1; .
 DR PIR; T14271; T14271.
 DR HSP; P16109; 1FSB.
 DR MGD; MGI:96083; Nr01.
 DR MGD; MGI:2447063; Od24.

DR InterPro; IPR006209; EGF_like.
 DR InterPro; IPR006210; IEGF.
 DR InterPro; IPR001258; NHL.
 DR InterPro; IPR006530; YD.
 DR Pfam; PF00008; EGF_3.
 DR Pfam; PF01436; NHL; 6.
 DR Pfam; PF05593; RNS_repeat; 5.
 DR SMART; SM00181; EGF; 6.
 DR TIGRFAMs; TIGR01643; YD_repeat_2x; 7.
 DR PROSITE; PS00022; EGF_1; 8.
 DR PROSITE; PS01186; EGF_2; 7.
 KW EGF-like domain.
 SQ SEQUENCE 2825 AA; 313483 MW; A120D98080686032 CRC64;
 Query Match 95.8%; Score 14269; DB 11; Length 2825;
 Best Local Similarity 93.4%; Pred. No. 0;
 Matches 2654; Conservative 47; Mismatches 52; Indels 88; Gaps 4;
 QY 1 MDVKEKPYRLTTRRDAERYTSSADSEBKAPQKSYSSSETLKAYDQDARLAYGSRV 60
 Db 1 MDVKEKPYRLTTRRDAERYTSSADSEBKAPQKSYSSSETLKAYDQDARLAYGSRV 60
 QY 61 KDVPQEAEEFCRTGANFTLRELGLLEEVTPPHGLYRTDGLPQCGYSMGAGSDADMEAD 120
 Db 61 KDVPQEAEEFCRTGANFTLRELGLLEEVTPPHGLYRTDGLPQCGYSMGAGSDADMEAD 120
 QY 121 TVLSEHVRVLRGSTRSGRSSCLSSRANSNLTLDTEHENTET-----DHPGGL 170
 Db 121 TVLSEHVRVLRGSTRSGRSSCLSSRANSNLTLDTEHENTET-----DHPGGL 170
 QY 165 -----DHPGGL 170
 Db 165 -----DHPGGL 170
 QY 171 QNHARLTPPPPLSHAHTPNQHHAAASINLNKRGFTPRSNPSAPTDSLSGEPPAGGAQ 230
 Db 241 QNHARLTPPPPLSHAHTPNQHHAAASINLNKRGFTPRSNPSAPTDSLSGEPPAGGAQ 300
 QY 231 EPHAQENWLLNSNIPLETRNLGKPPFLGTQDNLNLEMDILGASRHDGAYDGHFLFKPG 290
 Db 301 EPHAQENWLLNSNIPLETRNLGKPPFLGTQDNLNLEMDILGASRHDGAYDGHFLFKPG 360
 QY 291 GTSPLFTTSPGYPLTSTVYSPPPPLPRSTPRSTPRSTPRSTPRSTPRSTPRSTPRST 350
 Db 361 GTSPLFTTSPGYPLTSTVYSPPPPLPRSTPRSTPRSTPRSTPRSTPRSTPRSTPRST 420
 QY 351 ATLVIILAYFVAMHLFGLNWHLOPMEG--QVVEITEDTASSWPVPTDVLSPSGTGLT 408
 Db 421 ATLVIILAYFVAMHLFGLNWHLOPMEG--QVVEITEDTASSWPVPTDVLSPSGTGLT 480
 QY 409 PDRKGGTTEGKPSFFPDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLKENV 468
 Db 481 PDRKGGTTEGKPSFFPDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIIDHPVHLKENV 540
 QY 469 SLGKAALVGIYGRKGLPPSHQTFQDFVELLDGRRLLTQEARSLGEGTPRSGRTVPSSSHET 528
 Db 541 SLGKAALVGIYGRKGLPPSHQTFQDFVELLDGRRLLTQEARSLGEGTPRSGRTVPSSSHET 600
 QY 529 GFIOYLDGSIWHLAFYNDGKESEVVSFLTTAIESVDNCPNCGYNGDCISGTCCHQFLGFL 588
 Db 601 GFIOYLDGSIWHLAFYNDGKESEVVSFLTTAIESVDNCPNCGYNGDCISGTCCHQFLGFL 660
 QY 589 GPDGGRASCPVLCGNGQYMKGRCLCHGSKGAECDVPTNQCIDVACSNHGTCTGTCTIC 648
 Db 661 GPDGGRASCPVLCGNGQYMKGRCLCHGSKGAECDVPTNQCIDVACSNHGTCTGTCTIC 720
 QY 649 NPGYKGESECEVDNCPNCGYNGDCISGTCCHQFLGFLGFLGFLGFLGFLGFLGFLGFL 708
 Db 721 NPGYKGESECEVDNCPNCGYNGDCISGTCCHQFLGFLGFLGFLGFLGFLGFLGFL 780
 QY 709 DTGLSCDPSWTGHDCSIEICAAADCGHGVGVGGTCRCEGDMGMAACDQACHPRCAEHG 768
 Db 781 DTGLSCDPSWTGHDCSIEICAAADCGHGVGVGGTCRCEGDMGMAACDQACHPRCAEHG 840

QY	769	TCRDGKCECPGNGEHCCTTAHYLDVRVKBGCPGLNGNGRCRTLDLNGWHCVCOLGWRGA	828
DB	841	TCRDGKCECPGNGEHCCTT-----BGCPLCNGNGRCRTLDLNGWHCVCOLGWRGT	891
QY	829	GCDSMETACGSKNDNDGDLVDCMDPDCCLQPLCHINPLCLGSPNPLDIIQETQVPVQ	888
DB	892	GCDSMETGCGDKNDNDGDLVDCMDPDCCLQPLCHVNPCLGSPDLDIIQETQVPVQ	951
QY	889	QNLHSPYDRIKFLVGRDSTHIIPENPFDCGHACVIRGQVMTSDGTFPLGVNISFVNNPL	948
DB	952	QNLNSFYDRIKFLVGRDSTHSIPENPFDCGHACVIRGQVMTSDGTFPLGVNISFINNPL	1011
QY	949	FGYTIHQDGSFOLVNTGGLSIIILREAPFIQEHITLWLPDRFPVMTIIMRHEENI	1008
DB	1012	FGYTIHQDGSFOLVNTGGLSIIILREAPFIQEHITLWLPDRFPVMTIIMRHEENI	1071
QY	1009	PSDLSNFARPNPVSPPLTSPASSCAEKGPIVPEIQALQEBISISGCKMRLSYLSST	1068
DB	1072	PSDLSNFARPNPVSPPLTSPASSCAEKGPIVPEIQALQEBIVIAGCKMRLSYLSST	1131
QY	1069	PGYKSVLRISLTHPTIPFNLMKVLHVAVEGRLEFRKWFAPAAAPDLSYFFIWDKTDVYNQV	1128
DB	1132	PGYKSVVRISLTHPTIPFNLMKVLHVAVEGRLEFRKWFAPAAAPDLSYFFIWDKTDVYNQV	1191
QY	1129	FGISEAFVSGYEVESCPDLILWEKRTTVLQGYEIDASKLGWSLDKHALLNIQSGILHK	1188
DB	1192	FGISEAFVSGYEVESCPDLILWEKRTTVLQGYEIDASKLGWSLDKHALLNIQSGILHK	1251
QY	1189	GNENQVSOQPPIVSGIMNGRRRISCPSCNGLADGNKLLAPVALTCGSDGSLYVGF	1248
DB	1252	GNENQVSOQPPIVSGIMNGRRRISCPSCNGLADGNKLLAPVALTCGSDGSLYVGF	1311
QY	1249	NYIRIPFSGNVNILELRNKDFHSHSPAHKYLLATDPMGAVFLSDNSRRVKIKST	1308
DB	1312	NYIRIPFSGNVNILEM-----SHSPAHKYLLATDPMGAVFLSDNSRRVKIKST	1364
QY	1309	VYVKDLVKNSEVAVAGTGDQCLPDDTRCGDGGKATEATLTPRGITVDFKGLIYFVDGTM	1368
DB	1365	TVVKDLVKNSEVAVAGTGDQCLPDDTRCGDGGKATEATLTPRGITVDFKGLIYFVDGTM	1424
QY	1369	IRRIDONGIISTLGLNDLTSARPLSCDSVMDISQVLEWPTDLAINPMDNSLYVLDDNV	1428
DB	1425	IRRIDONGIISTLGLNDLTSARPLSCDSVMDISQVLEWPTDLAINPMDNSLYVLDDNV	1484
QY	1429	VLOISENHQVRIIVAGRPVHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYTAETDE	1488
DB	1485	VLOISENHQVRIIVAGRPVHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYTAETDE	1544
QY	1489	KKINRIQVTTSGEISLVAGAPSGCKNDANDCDFSGDDGVAKAKUNTPSSLAVCADG	1548
DB	1545	KKINRIQVTTSGEISLVAGAPSGCKNDANDCDFSGDDGVAKAKUNTPSSLAVCADG	1604
QY	1549	ELVAVDLGNIRIFIRKKNPFLNTQMYELSSPIDQELYLFTTGKHLVTSQSLPTGDIY	1608
DB	1605	ELVAVDLGNIRIFIRKKNPFLNTQMYELSSPIDQELYLFTTGKHLVTSQSLPTGDIY	1664
QY	1609	NFTYTGSDITLITDNNGNVNRDSTGMPVLVVPDQGVVMTGNTSALKSVTTQGH	1668
DB	1665	NFTYTGSDITLITDNNGNVNRDSTGMPVLVVPDQGVVMTGNTSALKSVTTQGH	1724
QY	1669	ELAMTYHGNSGLATKSNENGWTTTEYEDSGRLTNVTFTGOVSSFRSDTDSVHVQV	1728
DB	1725	ELAMTYHGNSGLATKSNENGWTTTEYEDSGRLTNVTFTGOVSSFRSDTDSVHVQV	1784
QY	1729	ETSSKDDVTITTNLSAGAFYTLLOQVVRNSYVIGADGSLRLLLANGMEVALQTEPHLLA	1788
DB	1785	ETSSKDDVTITTNLSAGAFYTLLOQVVRNSYVIGADGSLRLLLANGMEVALQTEPHLLA	1844
QY	1789	GTVPNPTVGRNVTLPTDNGNLNVWRQKQVQVTFEGRILRVHNRNLLSLDPRVTR	1848
DB	1845	GTVPNPTVGRNVTLPTDNGNLNVWRQKQVQVTFEGRILRVHNRNLLSLDPRVTR	1904

QY	1849	TEKIYDHRKFTLRILYDQAGRPSLWSPSSRLNGVNTYSPGGYIAGIQGIMSERMEYD	1908
DB	1905	TEKIYDHRKFTLRILYDQAGRPSLWSPSSRLNGVNTYSPGGYIAGIQGIMSERMEYD	1964
QY	1909	QAGRITSRIPADGKTWSYTYLEKSMVLLHSORQYIPEFDKNDRLSVWTPNVARQLET	1968
DB	1965	QAGRITSRIPADGKTWSYTYLEKSMVLLHSORQYIPEFDKNDRLSVWTPNVARQLET	2024
QY	1969	IRSVGYVRNIYQPEEGNASVIQDFTEDGHLHTPYLTGTRRVYKYOKLSKLAEITLYDTT	2028
DB	2025	IRSVGYVRNIYQPEEGNASVIQDFTEDGHLHTPYLTGTRRVYKYOKLSKLAEITLYDTT	2084
QY	2029	KVSTYDETACMLKTNLQNEGFTCTIRYQIGPLIDROJFRFRTEEGVNVARFDYNDNS	2088
DB	2085	KVSTYDETACMLKTNLQNEGFTCTIRYQIGPLIDROJFRFRTEEGVNVARFDYNDNS	2144
QY	2089	FRVTSMAVINETPLPIDLYRDDVSGTKTQFGKFGVIYDINQIITTAVMTHSKHFDAY	2148
DB	2145	FRVTSMAVINETPLPIDLYRDDVSGTKTQFGKFGVIYDINQIITTAVMTHSKHFDAY	2204
QY	2149	GRMKEVOYEIPRSLMYMTVOYDNMGRVVKELKVGYPANTTRYSYEYDADGQLOTYSIN	2208
DB	2205	GRMKEVOYETFRSLMYMTVOYDNMGRVVKELKVGYPANTTRYSYEYDADGQLOTYSIN	2264
QY	2209	DKPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRLGDVQYKMDDEGFLRQGGDI	2268
DB	2265	DKPLWRYSYDLNGLNHLSPGNSARLTPLRYDIRDRITRLGDVQYKMDDEGFLRQGGDI	2324
QY	2269	FEYNAGLLIKAYNRAGSWVRVYRGLGRRVSKSSHHLQFFVADLTNPVKVTHLYN	2328
DB	2325	FEYNAGLLIKAYNRAGSWVRVYRGLGRRVSKSSHHLQFFVADLTNPVKVTHLYN	2384
QY	2329	HSSEIITSLYDQGHUFAMELSSGDDEFFYACDNIGTFLAVFSGTGLMIKQIILYTAIGE	2388
DB	2385	HSSEIITSLYDQGHUFAMELSSGDDEFFYACDNIGTFLAVFSGTGLMIKQIILYTAIGE	2444
QY	2389	YMDTNPNFQIIIGYHGLYDPLTKLVHMGRRDYDVLAGRWTSPDHELWKLSSSNVMPFN	2448
DB	2445	YMDTNPNFQIIIGYHGLYDPLTKLVHMGRRDYDVLAGRWTSPDHELWKLSSSNVMPFN	2504
QY	2449	LYMEKNNPISNSODIKCFMTDVNSMLLTGFGOLHNVIPCYKPEDMDAMEPSYELIHTQM	2508
DB	2505	LYMEKNNPISNSODIKCFMTDVNSMLLTGFGOLHNVIPCYKPEDMDAMEPSYELIHTQM	2564
QY	2509	KTOEWDNSKSLGVQCEVQKAFVTLERFDQLYGSTIITSCQAPETKFKFASGSGVFGK	2568
DB	2565	KTOEWDNSKSLGVQCEVQKAFVTLERFDQLYGSTIITSCQAPETKFKFASGSGVFGK	2624
QY	2569	GVKFALXGDRVTTDIIISVANEDGRRVAAIILNHAHYLENLHFTIDGVDVTHYFVKPGPSEGD	2628
DB	2625	GVKFALXGDRVTTDIIISVANEDGRRVAAIILNHAHYLENLHFTIDGVDVTHYFVKPGPSEGD	2684
QY	2629	LAILGLSGRRTLENGVNVTVSQTINVTNGRTTTRYTDIOLQYCALCLNTRYGTTLDDEKA	2688
DB	2685	LAILGLSGRRTLENGVNVTVSQTINVTNGRTTTRYTDIOLQYCALCLNTRYGTTLDDEKA	2744
QY	2689	RVLELAPQAVRQAWAREOQRLREGEGBERAWTEGSKQVLTGRVQGVDPGFVIVEQY	2748
DB	2745	RVLELAPQAVRQAWAREOQRLREGEGBERAWTEGSKQVLTGRVQGVDPGFVIVEQY	2804
QY	2749	PELSDSANNIHEVROQSEWGR 2769	
DB	2805	PELSDSANNIHEVROQSEWGR 2825	

RESULT 3

Q9W7R3

ID Q9W7R3

AC Q9W7R3

DT 01-NOV-1999 (TREMBlrel. 12, Created)

DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE Ten-m4.


```
QY 1547 DGELIYVADLGNIRIFTRKKNKPFINTQNMVELSPIDQBLYLFDTTGCKHLYTQSLPTGDY 1606
Db 1602 DGELFIADLGNIRIRYVRNRKAFINPLNMVEISPIDDELYLFDVNASHVFTQSLTTGDY 1661
QY 1607 LYNFTYTGDDGDI TLITDNNGMVMNVRDSTGMPLWLVPPDGVVYWTMGNSALKSVTTQ 1666
Db 1662 LYNFTYSGBDLSISITDKXNRVIRRDSTGLPLWLGPDGQFTFTMGNTNALKSVAQ 1721
QY 1667 GHELAMTYHNGSGLLATKSNENGWTFYFDYSGRLTNVTFPTQVGSFRSDTSSVHV 1726
Db 1722 GQELAVTYHGGSGLLATKSNEDGWSFYEDYNGRLTNVTFPTGRVSVYRSDSDTVRV 1781
QY 1727 QVETSSKDDVTITNLSASAFYTLLODOVRNVSYYGAGSLRLLLANGMEVALQTEPHL 1786
Db 1782 QTEGSKEDITVTNLSASGTFYTLMDQOVKNYSYYIGLDSLRVLVLANGNEVSLHTEPHL 1841
QY 1787 LAGTVNPTVGRNVTLPIIDNGLNLVWRQKEQARGQVTVFGRRLRVHNRNLLSLDFRV 1846
Db 1842 LSGTVNPTISKRVNTPIDNGLNLVWRQKEQARGQVTVFGRRLRVHNRNLLSMDFRV 1901
QY 1847 TREKIYDDHRKFTLRILYQAGRPSLWSPSSRLNGVNTYSPGGYIAGIORGINSEME 1906
Db 1902 TRTEKVYDDHRKFTLRIRHYDHACRPTLWAPSSRLNGVNTYSPGGYIAGIORGTVSRME 1961
QY 1907 YDOAGRITSIFADGKTWSTYILEKSMVLLHROQVIFEFDKNDRLSSVTWPNVARQTL 1966
Db 1962 YDQGRITSIFADGKSNSTYILEKSMVLLLYSQRYIFEEDKNDRLSSVTWPNVARQTL 2021
QY 1967 ETIRSVGYXNIIYOPPEGNASVQDFTEDGHLHTYFLGTRVIVYKGLSKLAETLYD 2026
Db 2022 ETIRSIGYXNIVRPPENATLQDYSEDGLLQTLHQGTGRVIVYKGLSKRLLEILYD 2081
QY 2027 TTKVSFYDTAGMLKTINLQNSGFTCTIRYROIPLIDRQIRPFTTEEGMVARPDYND 2086
Db 2082 TTRAFSDYSAGMLKTIVGQSQFACITIRYROIPLIDRQIRPFTTEEGMVARPDYND 2141
QY 2087 NSFRVTSQMAVINETPLDLYRYDDVSGKTEQFGKFGVIYDINQIITAVMTHTKHPD 2146
Db 2142 NSFRVTSQMAVINETPLDLYRYDDVSGKTEQFGKFGVIYDINQIITAVMTHTKHPD 2201
QY 2147 AYGRMKEVOYEIRSLNMYWTVQYDNMGRVVKELKVGYPYANTRYSYEYDADGQLOTVS 2206
Db 2202 AYGRMKEVOYEIRSLNMYWTVQYDNMGRVVKELKVGYPYANTRYSYEYDADGQLOTVS 2261
QY 2207 INDKPLWRYSDYDNGNLHLLSPGNSAELTPLRYDIRDITRLGDVQYKMDDEGFLRQSG 2266
Db 2262 INDKPLWRYSDYDNGNLHLLSPGNSAELTPLRYDIRDITRLGDVQYKMDDEGFLRQSG 2321
QY 2267 DIFEYNSAGLLIKAYNPAGSNVRYRVDGLGRVSSKSHSHLOFFYADLNPYKVTHL 2326
Db 2322 DFEYNSAGLLIKAYNPAGSNVRYRVDGLGRVSSKSHSHLOFFYADLNPYKVTHL 2381
QY 2327 YNHSSEITSLYDLOQLHFAELSSGDEFYIACDNIGTFLAVFSGTGLMIKILYTAG 2386
Db 2382 YNHSSEITSLYDLOQLHFAELSSGDEFYIACDNIGTFLAVFSGTGLMIKILYTAG 2441
QY 2387 EYMDTNPFPQIIGYHGGLYDPLTKLVHNGRDRYDVLAGRWTSPDHLWKHLSGNVNP 2446
Db 2442 EYLDNSPSPQVIGYQGGLYEPLTKLVHNGRDRYDVLAGRWTSPDHLWKHLSGNVNP 2501
QY 2447 FNLVYFKNPNPI NSODIKCFMTDVNSWLLTFQGLHNVTIPGYKPDMDAMPSEYELIHT 2506
Db 2502 FNLVYFKNPNPI NSODIKCFMTDVNSWLLTFQGLHNVTIPGYKPDMDAMPSEYELIHT 2561
QY 2507 QMKTQWDSKSIIGVQCEVQKLFVTLEREDQLYGSITTSQOQAPKTKFASGSGVF 2566
Db 2562 QIKTQWDSKSIIGVQCEVQKLFVTLEREDQLYGSITTSQOQAPKTKFASGSGVF 2621
QY 2567 KGKVKALKDGRVTTDIISVANGEDGRVAAIINHAHYLENHLHETIDCVTHYFVKGPSE 2626
Db 2622 KGKVKALKDGRVTTDIISVANGEDGRVAAIINHAHYLENHLHETIDCVTHYFVKGPSE 2681
QY 2627 GDAILGLSGRRGLENGVNVTVSQINTVLNGRTRRYTDIQLQYALCLNTRYGTTLDDEE 2686
```

```
Db 2692 GDLULLGTVGQRTLETGVNVTVSQVNNVLGGRSRRITDQMQIGTULSNVRGSSVDEE 2741
QY 2697 KARVLELARQAVRAQWARQORLRGEGLRAWTEGKQQLVSTGRVQGYDGFVVISVE 2746
Db 2742 KRVVLELARQAVATAWAHERHRLRQGEESRAWTDGERQQLSSGRVQGYEGFYIVSD 2801
QY 2747 QYPELSDSANNIHPMRQSEMGR 2769
Db 2802 QPELTNDINNVPWFQTEMGR 2824

RESULT 4
QWTS6 PRELIMINARY; PRT: 2715 AA.
ID QWTS6;
AC QWTS6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Ten-m3
GN OD23 OR TEN-M3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c; TISSUE=Brain;
RA Coshishi T., Zhou X., Feng K., Richter B., Moergelin M., Perez M.T.,
RA Su W., Chiquet-Ehrismann R., Rauch U., Faessler R.;
RT "Mouse Ten-m3 is a new family of dimeric type II transmembrane
RT proteins expressed in many tissues.";
RL J. Cell Biol. 0:0-0(1999).
DR EMBL; AB025412; BAA77398.1; -
DR HSPF; P01492; 1VNB.
DR MGD; MGI:1345183; Odz3.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0004623; F:phospholipase A2 activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0016042; P:lipid catabolism; IEA.
DR InterPro; IP006209; EGF like.
DR InterPro; IP006210; IEGF.
DR InterPro; IP002049; Laminin_EGF.
DR InterPro; IP001258; NHL.
DR InterPro; IP001211; PhospholipaseA2.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF01436; NHL; 6.
DR Pfam; PF05593; RNS_repeat; 6.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00181; EGF; 6.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 6.
DR PROSITE; PS00022; EGF_1; 8.
DR PROSITE; PS01186; EGF_2; 7.
KW EGF-like domain.
SQ SEQUENCE 2715 AA; 303062 MW; 598F46A77334C2E1 CRC64;

Query Match 68.7%; Score 10227; DB 11; Length 2715;
Best Local Similarity 66.7%; Pred. No. 0;
Matches 1864; Conservative 370; Mismatches 457; Indels 102; Gaps 19;

QY 1 MDVKERYSLTR-RSDAERRYTSSADSEEGKAP-QKSYSSSETLKAYDQD-ARLAYG 57
Db 1 MDVKERYPSLTKSRREKERRYTNSSADNEECRVTKQKYSSETLKAFDHYRLLYG 60
QY 58 SRVKDIPQEAEEFCRTGANTFRELGLVEEYTPPHGLTYRTDGLPGCGSMGAGSDAM 117
Db 61 NRVKDLVHREADEYTRQGNFTLRQLGVCSATERGVAFCAEMGLPHRGYSISAGSDAT 120
QY 118 BADTVLSPHPVRLWGRSTRSGSSCLSSRANSNLITLTDTEHEN--TETCHPGGLQNH 174
Db 121 ENEAVMSPPEHAMRLWGRVSGSSCLSSRNSALTITDTEHENRSDSESEQSNPNQP 180
```



```

Db 2277 HVYHSSSEITSLYYDLQGLHAMEISSGDEFFIASDNTGTPLAVFSSNGLMLKQTYTA 2336
Qy 2385 YGEIYMDTNFNFIIGYHGLYDPLTKLVHMGRRDYDLVLAGRWTPSDHMLKXHLSSNV 2444
Db 2337 YGEIYFDSNVDQAVIGFGLYDPLTKLHFGERYDYLILAGRWTPDI-EIWKRI-GKDP 2395
Qy 2445 MPFNLYMKNPNISNODIKCFWTDVNSMLLFFGOLHNVIYCYKPKPMDMAMEPSYELI 2504
Db 2396 APFNLYMFRNNPASKIHDVKDYITDVNSMLVTFGHLNVAIFGPVPKFDLTERPSYELV 2455
Qy 2505 HTQMKTEWNSKSIILGVQCEVQKQKAFVTLERFDLPOLYGSTITSQQAPKTKK----- 2558
Db 2456 ----KSQWEDVPPIFVQOQVAFQAKAFLSLGMAEV-----QVSRKAGAEQSW 2502
Qy 2559 --PASSGSVFGKVKFALKDGRVTTDIISVANEDGRRVAALNHAHYLENLHFTIDGVT 2616
Db 2503 LWFATVSLKLGKWLAVSGRVQTNVNTANEDCIVKAAVLNNAFYLENLHFTIEGKT 2562
Qy 2617 HYFVKPSPSGDLAILGLSGRRLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLN 2676
Db 2563 HYPIKTTTPESDLTLTSGKALENGINVTVSQSTTVNGRTRRFADVEMQFGALALH 2622
Qy 2677 TRYGTTLDEKARVLELQARAVQAWAREQOQLREGESGLRAWTEGEKQOVLSTGRVQ 2736
Db 2623 VRGMTLDEKARLEQARQARALARAWAREQQVRDGEAGRLWTEGEKQQLSAGKVG 2682
Qy 2737 YDGPFFVISVQYPELSANNIHFMRQSEMR 2769
Db 2683 YDGYVLSVQYPELADSANNIQLRQSEIGKR 2715

RESULT 5
Q804R1 PRELIMINARY; PRT; 2192 AA.
AC Q804R1; 2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 25, Last annotation update)
DE SI:zC197811.1 (odd Oz/ten-m homolog 4) (Fragment).
GN OD24.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Babbage A.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL307754; CAB61260.1; -
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR006530; YD.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF01436; NHL; 4.
DR SMART; PRO0011; EGF-LAMININ.
DR SMART; SM00181; EGF; 8.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 5.
DR PROSITE; PS00022; EGF_17_8.
DR PROSITE; PS01186; EGF-2; 7.
FT NON TER.
SQ SEQUENCE 2192 AA; 243920 MW; 063A4B36DA3718A5 CRC64;

Query Match 67.0%; Score 9968; DB 13; Length 2192;
Best Local Similarity 80.3%; Pred. No. 0;
Matches 1774; Conservative 242; Mismatches 176; Indels 16; Gaps 2;

Qy 562 SVNCPSCNCGNDCTSGTCHCFGLGTPDCGRASCPLVCSNGQYMKGRCLCHSGWKGA 621
Db 1 SIDDPCSNCFNGSDVSGNCHCFGRBPDCSRASCPLVCSNGQYMKGRCHSGWKGS 60

```

```

Qy 622 ECDVPTNOCIDVACSNHCTCITGTCICNPGYKGBSCSEVDCMDPTCSGRGVGVGRGECHCF 681
Db 61 ECDVPTNOCIDITCSGHGTCIVGTCICNPGYKGBSCSEVDCMDPTCSGRGVGVGRGECHCF 120
Qy 682 VGMGTCNETPRATCLQCSGHGTFLLPDTGLCSDDPSWTGHDSEIEICAAACDGGHGVCG 741
Db 121 VGMGPGCESPRASCMQCSGHGSLADTNTCNCDHNWTHGDCSTELCAADCGHGVCA 180
Qy 742 GTCRCEDGMAACDQACHPRCAEHGTCDKCECSCPGWNGEHCTTAHYLDRVKEGCP 801
Db 181 GSCRCDEGMMGTGCEQACHPRCSEHGTCKDKECSCPGWNGEHCTT-----EGCP 231
Qy 802 GLCNGNGRCTLIDLNGHCVCOQGWGAGCDTSMETACDGDNDGDLVDCMDPDCCLOP 861
Db 232 GLCNGNGRCTLGNNGWYCVQQLGWRGAGCDTSMETACDGDNDGDLTDCMDPDCCLOA 291
Qy 862 LCHINPLCLGSPNPLDIIQETQVPVVSQONLHFSFYDRIKFLVGRDSTHIIPEGNPFDDGHA 921
Db 232 SCHTTSLCVGSPLDIIQETQISSLSLTQSFYQRIHFLVGRDSTHVI PGVNPFDGIHA 351
Qy 922 CVIRGQVMTSGTPLVGNISFVNPLFGYTIISRODGSFDLVNTGTSIIILRTERAPFIT 981
Db 352 CVIRGQVMTSGTPLVGNISFVNPLFGYTIISRODGSFDLVNTGTSIIILRTERAPFIT 411
Qy 982 QEHTLWLPWDRFFYMETIIMRHEENEIPSCDLSNFAFPNPNVSPSPULTSPASSCAEKGP 1041
Db 412 QEHTLWLPWGRFFVWMDTIIMRHEVNDIPSCDLSSTFRMPILVPAFLTAFACTPERGIV 471
Qy 1042 VPEIQALQEEISISGCKMRLSYLSRTPGYKSVLRISLTHPTIPFNLKMLVAVVEGRL 1101
Db 472 VPEIQALQEEISISGCKMRLSYLSRTPGYKSVLRISLTHPTIPFNLKMLVAVVEGRL 531
Qy 1102 FRKWFAPADLSYFYIWDKTDVYNQKVFGLSEAFVSVGYEYESCPDLILWEKTTVLQGY 1161
Db 532 FRKWFAPADLSYFYIWDKTDVYNQKVFGLSEAFVSVGYEYESCPDLILWEKTTVLQGY 591
Qy 1162 BIDASKLGGWSLDKHALNTQSGILHKGNGENQFVSQOQPPVIGSIMGNGRRRISCPSCN 1221
Db 592 ETTASNLGGWSVDKHALNTQSGILHKGNGENFISQOQPPVIGSIMGNGRRRISCPSCN 651
Qy 1222 GLADGNKLLAPVALTCGSDGSLVYVDGPNYIRRRIPPSGNVTNILELRNKDPRHSHSPAHKY 1281
Db 652 GLADGNKLLAPVALTCGSDGSLVYVDGPNYIRRRIPPSGNVTNILELRNKDPRHSHSPAHKY 704
Qy 1282 YLATDPMGSAVFLSDNSRRVFKTKSTVVVKDLVNSEVAVAGTQDQCLPDDTRCGDGGK 1341
Db 705 YLATSPVGMWLYSDTSRKFVKSILYAVKDAKLELVAGTQDQCLPYDETRCGDGGK 764
Qy 1342 ATEATLNPRTGIVDYKEGLYFVYDGTMRIRIDONGIISTLLGNDLTASPLSCDSVMDI 1401
Db 765 AVEATLNPRTGIVDYKEGLYFVYDGTMRIRIDONGIISTLLGNDLTASPLSCDSVMDI 824
Qy 1402 SQVLEWPTDLAIMPNDNSLYLDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLSKV 1461
Db 825 SQVLEWPTDLAIMPNDNSLYLDNNVVLQISENHQVRIVAGRPMHCQVPGIDHFLSKV 884
Qy 1462 AIHATLESATALAVSHNGVLYIAETDEKKINRQVTTSGEISLVAGAPSGCCCKNDANC 1521
Db 885 AIHATLESATALAVSHNGVLYIAETDEKKINRQVTTSGEISLVAGAPSGCCCKNDANC 944
Qy 1522 DCFSGDDGYAKDAKLNTPSSILAVCADGELYVADLGNIRIRIFIRKKNPFLNTQNYELSSP 1581
Db 945 DCYSDDGYAKDAKLNTPSSILAVCADGELYVADLGNIRIRIFIRKKNPFLNTQNYELSSP 1004
Qy 1582 IDQELYLFDTTGKHLTYQSILPTGDIYNFTYTGDDITLITDNGNMMVNRVRRSTGMPLW 1641
Db 1005 IDDELYLFDVNASHVFTQSILPTGDIYNFTYTGDDITLITDNGNMMVNRVRRSTGMPLW 1064
Qy 1642 LVVPDGOVYVMTNTSALKSVTTQGHLEAMTYVHNSGLLTKNSNENGWTTTFYEDVSFG 1701
Db 1065 LMGPDDGQTFMTGTNNALSAVAQGEIIVMTYHSGSGLLTKNSNENGWTTTFYEDVSFG 1764

```



```
QY 1702 RLTVPTPTGOVSSFRSDTSSVHVVOVETSSKODVTITTNLSASGABYVTLLOQOVNSYY 1761
Db 1125 RLTVPTPTGVSSYRSDSTSVRQTEGSKBDITVTNLSASGTFTYLMQOVKNSSYY 1184
QY 1762 IGADGSLRLILANGMEVALQTEPHLLAGTVNPTVGKENVTLPIPDINGLNLVEMVQRKEQAR 1821
Db 1185 IGLDGLSLRLVLANGMEVSLTEPHLLSGTVNPTISKENVTLPIPDINGLNLVEMVQRKEQAR 1244
QY 1822 QGVTVFGRRLVHRNLLSLDPDRVTEKTYDDHRKFTLRILLYDQAGRSLWSPSRLN 1881
Db 1845 QGVTVYGRRLVHRNLLSLDPRVTEKTYDDHRKFTLRILLYDQAGRSLWSPSRLN 1304
QY 1882 GWNVTYSPGGYIAGIQIGIMSERMEYDOAGRITSRIFADGKTWSTYTYLEKSMVLLHLSOR 1941
Db 1305 GWNVTYSPGGHIAIQIGTMSVRMEYDQNGRITSKIFADGKSWSTYTYLEKSMVLLYSOR 1364
QY 1942 QYIFEFKNDRLSLSVTPNVARQLETIRSVGYRNTYQPEGNASVIQFTEDGHLHHT 2001
Db 1365 QYIFEFKNDRLSLSVTPNVARQLETIRSVGYRNTYQPEGNATVLOQSGDGLLQ 1424
QY 2002 FYLGTGRVLYKYGKSLKSLATLTDYTKVSTFYDETAKMLKTINLQNEGTCTIRYQIG 2061
Db 1425 IHQGTGRVLYKYGKSLKSLLEILYDTRIAFSYDESAGMLKTIVLQSGEGFACIRYQIG 1484
QY 2062 PLIDRQIFRFEPEGMVNAFEDYNDNSFRVTSQAVINETPLPIDLYRYDDVSGKTEQFG 2121
Db 1485 PLIDRQIFRFESEGMVNAFEDYNDNSFRVTSQAVINETPLPIDLYRYDDVSGKTEQFG 1544
QY 2122 KFGVIYDINQIITAVNTWTHKHPDAYCRMEKVOYEIFRSLMYMTVQYDNGGRVVKEL 2181
Db 1545 KFGVIYDINQIITAVNTWTHKHPDAYCRVKEVOYEIFRSLMYMMVQYDNGGRVAKEL 1604
QY 2182 KVGPYANTTRYSEYDADGQLOQTVSINDKPLWRSYDLNGLHLLSPGNSARLTPLRYDI 2241
Db 1605 KVGPYANTTRYAYEYDADGQLOQTVSINDKPLWRSYDLNGLHLLSPGNSARLTPLRYDI 1664
QY 2242 RDRITRLGDVOYKMDGDFLQRQCGDIFEVNSAGLLIKAYNRAGSVVRVYDGLGRVRS 2301
Db 1665 RDRITRLGDVQYRDEDFLQRQCGDFFEVNSAGLLIKAYNRAGSVVRVYDGLGRVRS 1724
QY 2302 SKSHSHHLOPFYADLNPRTKVTLYNHSSEITSLYDLOQHLFAMELSSGDEFFIACD 2361
Db 1725 SRSTQGHHLQPFYADLNPRTKVTLYNHSSEITSLYDLOQHLFAMELSSGDEFFIACD 1784
QY 2362 NIGTPLAVFGTGLMIKQILYATXGEIYMDTNPNFQIIGYHGLYDPLTKLVHMGRRDY 2421
Db 1785 NIGTPLAVFGSAGLMIKQILHTAFGEVYLDNSNPSFQIVIGYQGGLYEPPLTKLVHMGRRDY 1844
QY 2422 DVLAGRNTSPDHELWKHLSNPNVFNLYMFKNNPISNSODIKCFMTDYNMILLTFGFG 2481
Db 1845 DVLAGRNTTPDHDLWKHLSNPNVFNLYMFKNNPISNSOETKCYMTDYNMILLTFGFG 1904
QY 2482 LHNVPYKPDMDAMBPSYELIHTOMKTOBWDNSKSLGVCQEVQKAKAFVTLERFDQ 2541
Db 1905 LYNVPGYKPEVTDAMBPSYELIHTQIKTOBWDSTKSVLGVQCEVQKAKAFVTLERFDQ 1964
QY 2542 LYGTITSCQAPKTKFASGSGVFGKGVKFPALKDGRVTTDITISVANEDGRRVAAILNHA 2601
Db 1965 IYSASDSGCPPTPLHTLTFATGTSFLFGKGVKVAIREGRVEADIISLANEDGRRVAAILDKA 2024
QY 2602 HYLENLHFTIDGVTHYFVKRPGPEGDLAILGLSGRRRTLENGVNVTVSGINTVLANGRTR 2661
Db 2025 SYLDQLHFTTAGLTHYFVKSGLVEGDLGLMTVGORTLETGVNVTVSGQNVVWLGRRSR 2084
QY 2662 RYTDIOLQYGALCINTRYGTTLDEKARVLELAFQRAVROAWAREQORLEGEGLRAWT 2721
Db 2085 RITDIOMQYGTLSLNVRYGSDVEKVRVLELAFQRAVATAWAHERHRLAQEGEGRAWT 2144
QY 2722 EGEKQVLSGRVQGYGDFVIVSVOYPELSDSANNHFHQSEMGR 2769
Db 2145 DGERQQLSSGRVQGYGDFVIVSVDQPFELADNINNHFJRQTEMGR 2192
```

```
RESULT 6
Q9W6V6 PRELIMINARY; PRT; 2705 AA.
AC Q9W6V6;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Teneurin-1.
GN TEN-1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
CX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
SEQUENCE FROM N.A.
RP TISSUE=Brain;
RX MEDLINE=99276585; PubMed=10341219;
RA Minet A.D., Rubin B.P., Tucker R.P., Baumgartner S.,
RA Chiquet-Ehrismann R.;
RT "Teneurin-1, a vertebrate homologue of the Drosophila pair-rule gene
RT ten-m, is a neuronal protein with a novel type of heparin-binding
RT domain.";
RL J. Cell Sci. 112:2019-2032(1999).
DR EMBL; AJ238613; CAB43098.1; -.
DR HSSP; P00750; 1TPG.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR001458; NHL.
DR InterPro; IPR006530; YD.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF01436; NHL; 5.
DR SMART; SM00181; RHS repeat; 6.
DR TIGRFS; TIGR01643; YD_repeat_2x; 7.
DR PROSITE; PS00022; EGF_1; 8.
DR PROSITE; PS01186; EGF_2; 7.
KW EGF-like domain.
SQ
SEQUENCE 2705 AA; 302385 MW; 230F03D1999037D2 CRC64;

Query Match 64.8%; Score 9642; DB 13; Length 2705;
Best Local Similarity 62.9%; Pred. No. 0;
Matches 1750; Conservative 405; Mismatches 535; Indels 94; Gaps 17;

QY 1 MDVKRKPYRSLTR-RDAERRYTSSSADSEBKAPQKYSSESSTLKAYDQARLAYG-- 57
Db 1 MEQMDCKPYQPLSKVKHEVDLTYTSSDESDGRKQROSYDSRETNLNEYSQELRLNYSQ 60
QY 58 SRVKDIVQSEAE--RFCRTGANFTLRGLGEVTPPH--GTLRYRTDI-GLPCCGYSMGNG 112
Db 61 SKRKNXTDQSTQDMFCET-----PHILCSGYQDLHGVSSEHSYPLEYVG 104
QY 113 SDADMEADTVLSPEHPVRLWGRSTRSGRSSCLSSRANGNLTLTDEHE---NTETDHGG 169
Db 105 SDVDTETEGGASPDHALRWWMGKSEHSSCLSSRANSALSLTDDHEKSDGENDMPGS 164
QY 170 LQNHARLR--TTPPPPLSHAHTPNQHAASINSNLNGNFTPRSNPSP-APTDSLSEPPA 226
Db 165 PHNQFTFRPLPPPPPPHACTCTRKPPPAADSLQRRSMTTRSQPSAAPT-----PP 216
QY 227 GGAQPAQAQENWLLNSNIPLETRNLGKQPFGLTIQDNLIEMDILGASRHDGAYSDGHFL 286
Db 217 TQTSQVHLHNSWLLNSNIPLETR-----HPL 243
QY 287 FKPG-GTSLPFCITSPGYPLTSSVYSPPPPLPRSTFARPAFNLLKPSKYCNWKAALS 345
Db 244 FKHGSGSAIFSAASQNYPLTSNTVYSPPPPLPRSTFARPAFNLLKPSKYCNWKAALS 303
QY 346 AVIVSATLVILLAYFVAMHLFGLNHLQPMQMEQVETEDTASSWPVTDVLSYSGGTG 405
Db 304 ATAITVLTALLLAYIVIAVHLFGLTQWQFVQGLYENGVSKNGKGAESTDTTYSYFGG-- 361
QY 406 LETPDRKKGKGTTEGKPSFFPEDSPDSIDVGRRASQKIPPGTFRASQVFIHPVHLK 465
```

Db 362 -----KVSDTEKK-----VFOKGRADTGEVEIGAQMOTIPPLGFWRFQITIHHPVYLK 412
QY FNVSLGKAAALVGIYGRKGLPSPHSTQDFVLLDGRLLTQEAESLSTPRQSGRTVPPSS 525
Db 413 FNI SLAKDLSLGIYGRNNIPPTHTQDFVKLMGKQLIKQEPKNSBEPOQAPRNLILTSL 472
QY HETCFIYLDSGIWHLAFYNDGKESSEVVSFLTIAESVDNCPNCYNGDCISGTCHEFL 585
Db 473 QETCFIEYMDGAMHAFYNDGKVEQVFLTTAIEVLDDCSTCNGNGBECISGHCHQFP 532
QY 586 GFLGPDGRASC PVLCSNGOYKGRCLCHSGWKGAECVPTQCTIDVACSNHGTCTIGT 545
Db 533 GFLGPDCAKDCSPVLCSENGEYKKGHCVRNGWKGPEDCPBEQCCTDPTCFGHGTIMGV 592
QY 646 CINCNGYKGSCEBVDQMDPTCSGRGVCVRGECHECFVGMGNTNCETPRATCLDQCSGHGT 705
Db 593 CICYPGYKGLCEBEDCLDPWCSGHGVCVCGECHCSAGNGVNCETSLPICQEHCSGHGT 652
QY 706 FLPTGLCSDPSPWTHGDCIEICADCGGHGVCVGTCTCECDGNGWGAACDQSAACHPCA 765
Db 653 FLDDVGLCSCEPQWGTGSDCTELCTLDGSGHGVCSRGICQCEBGGVGPCTCEERTCHSHCA 712
QY 766 BHGTCRGCKECSPGWNGEHCITIAHYLDRVVVKGSCGLNGNGRCRTLDLNGWHCVCOLGW 825
Db 713 EHGCKDKGCKECSPGWEGDHCTI-----DCCPGLCYGNRCRTLDQNGWHCVCOQGW 763
QY 826 RGACDTSMETACDSDKNDGDGLVDCMDPCCQLPCLHINPLCLGSPNPLDIIQETQVP 885
Db 764 SGSCNVMEMACGDNLDNDGDGLTDCVDPCCQONCYASPLCQSGPDPDLDIQHSQPP 823
QY 886 VSQONLHSFVDRKFLVGRSTHIIIPGENFPDGGHACVIRQVMTSDGTFPLVCVNSFVN 945
Db 824 FSQHPPLFVDRIRFLGKSTHVPDGISEFERRASVIRQVVAIDGTFPLVGVNVSFLH 883
QY 946 NPLFGYITISQDGSFDIYVNGGISIIILFERAPFIQEHFTLWLPDRFFMETIIMRHEE 1005
Db 884 HDEGYITISQDGSFDIYVNGGISIVLVFDESFFISEKRTLWLSWNRVFIYDKVMQRAE 943
QY 1006 NETPSCDLSNFPARNPVVSPSPITSFASCAEKGPIVPEIQALQBEISGCKORLSYLS 1065
Db 944 SDIPSCDVSSFISNPVLSPLTAFGSCSPERGTVIPELQVYQBEIPSPSSFVLSYLS 1003
QY 1066 SRTPGYKSLRISLTHPTIFNLKMYHMAVEGRFLRKFKAAPADLSYVFIWKTDDVYN 1125
Db 1004 SRTPGYKSLRVLTHPTIFSGWKVHLIIIAVEGRLLQKFFPAAANLVTFPAKNTDIYG 1063
QY 1126 QKVEGLSEAFVSVEYYESCPDILWEKRTTVLQGEYI DASKLGGWSLKDHHALNTQSGI 1185
Db 1064 QKVSGLAEAMVSGVEYETCFDPLWEKRTVILQGFEMDASNLGWSINKHHVNPQSGI 1123
QY 1186 LHKNGENOFVSQOPPVIGSIMNGRRBSISCSNGLADGNKLLAPVALTCTCSGSLYV 1245
Db 1124 VHKNGENMFISQOPPVISTMNGHORSVSCNGLAUNSLKLPAPVALTSGPDGSVYI 1183
QY 1246 GDFNYIRRIIPSGNVNTNILELRNKRDFRSHSPAKHYIYATDPMSGAVFLSDNSRRVFKI 1305
Db 1184 GDFNFVRIIPSGNSIGILELRNRDTRHSTS PAKHYIYAVDPVYESLSYLDNTTRAVYKA 1243
QY 1306 KSTVVVZOLVKNSEVAVAGTGDCLPDDTRCGGGKATEATLNPRTITVDKFLIYFVD 1365
Db 1244 KSLIETDALKANDVAVAGTGDCLPFDQSCGDKGKASEASLNSPRGITIDKHGFIYFVD 1303
QY 1366 GTMIRRIDQGIISTLGSNDLTSARPLSCDSVMDISQVLEWPTDLANPMDNSLYVLD 1425
Db 1304 GTMIRKIDENGMITIIIGNSLSTQPLSCDSGMDITQVLEWPTDLTNPLDLSLYVLD 1363
QY 1426 NNVLQJSENHQRIVRAGRMHCOVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAE 1485
Db 1364 NNIVLQJSESRVRRIIAGRIHCOVPGIDHFIYSKVAIHSTLESARAIIVSHSGIPIRE 1423
QY 1486 TDEKINRIQVTTSGEISLVAGAPSCDCKNDANCDPFGDGGYAKDAKLNTPPSLAVC 1545

Db 1424 TDERKINRIQVTTNGEISIIAGAPSCDCKIDPNCDFSGDGGYAKDAKLKAPSSIAVS 1483
QY 1546 ADGELYVADLGNIRIRPIRKNKPLFNTONMYELSPIDQELYLFDTTGKELYTOSLPTGD 1605
Db 1484 PDDTLYVADLGNIRIRAVRNKAHLSDTNMYEIASPADQELYOPTINGTHLHTLNLITRD 1543
QY 1606 YLYNFTYTGDDITLITDNNNGMNVNRDSTGMPLWLIVPDGQVYVMTGNTNSAKSVTT 1665
Db 1544 YLYNFTYSGEGDVATITSSNGNSVHIRRDTSGPLWVVPVGGQVYMLTISNGVLEKRYA 1603
QY 1666 OGHELANMTYHSGNSGLLATKSNENGWTFEYEDSGRLTNVTPPTGOVSPFRSDTSSVH 1725
Db 1604 QYINLALMTYPTGNTGLLATKSDENGWTFEYEDSGHLTNATPTGEVSSFSHDSVEKLTR 1663
QY 1726 VQVETSSDDVYITITNLSASGAFYLLQDQVRNSYIIGADGSLRLLLLANGVEVALQTEPH 1785
Db 1664 VELDTSNRENMTATNFSATSTIYTLKQDNTQNIYRSPDGLRVTFASGWEITLNTPEH 1723
QY 1786 LLAGTVNPTVGRNVTLPIDNLGNLVEWRQKEQARQVTVFGRBLRVHVNENLISLDPDR 1845
Db 1724 ILAGVVSFTLCKCNISLPGEHNSNLIEMWRQREQTGKNISTFERRLRAHNRNLLSIDFDH 1783
QY 1846 VTRTEKIYDDHRKFTLRILYDQAGRPSLWSPSSRLNGVNVITYSPGYIAGTQRCIMSERM 1905
Db 1784 VTRCKIYDDHRKFTLRILMYDQGRPVLWSPISKYNEVNIYSHSLVTVIQRGTWTEKM 1843
QY 1906 BYDQAGRITSRIFADGKTWSTYILEKSMVLLLSQROYIIFEDKNDRLSSVTMPNVARQT 1965
Db 1844 BYDPSGNIISRTWADGKTWSTYILEKSMVLLLSQROYIIFEDQSDYLLSPWSPMVYHA 1903
QY 1966 LETTRSGVYENIYQPPPEGNASVIQDFTEDGHLLHTFVLGTGRVVIYKGYKLSKLAETLY 2025
Db 1904 LQTMUSVGYENIYTPPSSGAFFIQDTRDGRLLQTLYPGTGRVLVLYKSKQSRLSILY 1963
QY 2026 DTKVSYFTYDTAGMLKXTINLQNEGFTCTIRYRQIGPLIDRQIRPFTTEGVMNARFVNY 2085
Db 1964 DTKVTFYESSGVIKTIHLMHDFCTIRYRQIGPLIDRQIRPFTTEGVMNARFVNY 2023
QY 2086 DNSFRVTSQMAVNETPLPIDLYVDVDSGTEQFGKGVYVYDINOIITANTWHTKHP 2145
Db 2024 -NFRVTSQMAVNETPLPIDLYRVDVDSGTEQFGKGVYVYDINOIITANTWHTKHP 2082
QY 2146 DAYGERMEVQVEIIFRSLMYMTVQYDNMGRVYKELVGYFANTRYSYEYDADQLOQTV 2205
Db 2083 SANGQVIEVQVEILKSIAYWMTIQYDNMGRVYKELVGYFANTRYSYEYDADQLOQTV 2142
QY 2206 SINDKPLWRYSDYDNGNHLSPGNSARLPLRVDIRDRITRLGDVQYKMDQDQFLQRG 2265
Db 2143 SVNDKTQWRYSDYDNGNHLSPGNSARLPLRVDIRDRITRLGEIYQYKMDQDQFLQRG 2202
QY 2266 GDIPEYNSAGLLIKAYNAGSWSVRYVDGLGRVSSKSSHSHLOFFYADLTNPTKVTH 2325
Db 2203 NEIPEYNSAGLLIKAYNAGSWSVRYVDGLGRVSSKSSHSHLOFFYADLTNPTKVTH 2262
QY 2326 LYNHSSSEITSLYDLOQHLFAMELSGDEBFYIACDNGITPLAVSFGTGLMKOILWYAY 2385
Db 2263 LYNHSSSEITSLYDLOQHLFAMELSGDEBFYIACDNGITPLAVSFGTGLMKOILWYAY 2322
QY 2386 GEIYMDTNPFIQIIGYVHGGYLDPLTKLVHMGRRDYPVLAGRMTSPDHELWKHLSNNVM 2445
Db 2323 GEIYQDITNPDPQVYVHGGYLDPLTKLVHMGRRDYPVLAGRMTSPDHELWKHLSNNVM 2381
QY 2446 PFNLYMFKNNNPISNSQDIKCFMTDVSNSWLLTFGQLHNVPYKPKPMDAMPESEYLIH 2505
Db 2382 PFNLYSPENNYPVGRIQDAKYITDIGSWLELFGQLHNVPYKPKPMDAMPESEYLIH 2441
QY 2506 TOMKTOEWDNSKSTILGYQCEVOKOLKAFVTLERFDQLYGSTITSQOAPKTKKFASSGSV 2565
Db 2442 LQTKTQEMDPOKTIILGQCELOKQLRNFIISLVQDPMTPRYSDGKYGEGVQKPRFAALPSV 2501
QY 2566 FGKGVKALKDGRVTTIISVANEDGRRVAAIILNHAHYLENLHFTIDGVTHYFVKPGPS 2625
Db 2502 FGKGIKFAIKDGIYVADIIGVANEDSRRIAILNNAHYLENLHFTISGRTHYFVKPGPS 2561

```

Qy 2626 EGDIAILGCGRTLENGVNVTVSQINTVLNGKTRRYTDIQLQYGALCLNTRYGTLDE 2685
Db 2562 EEDLSLNTGGRIELGVNVTVSQMTSVNGKTRRFADIQHQGALCFNVRYGTTVEE 2621
Qy 2686 EKARVLELARCRAVQRAWEQOQLREGEGLRAWTEGEKQOVLSTGRVQGYDGFVISV 2745
Db 2622 EKNHVLVAQRAVAQAWTKEORRLOEGEGIRAWTDGEKQQLNTRGVQGYDGFVISV 2681
Qy 2746 EQYPELSDSANNIHFMRQSEMGR 2769
Db 2682 EQYLELSDSANNIHFMRQSEIGR 2705

RESULT 7
Q9DERS
AC Q9DERS; PRELIMINARY; PRT; 2802 AA.
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Teneurin-2 (Fragment).
GN TEN2.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=20581705; PubMed=11146505;
RA Tucker R.P.; Chiquet-Ehrismann R.; Chevron M.; Martin D.; Hall R.J.;
RA Rubin B.P.;
RT "Teneurin-2 is expressed in tissues that regulate limb and somite
RT pattern formation and is induced in vitro and in situ by FGF8."
RL Dev. Dyn. 220:27-39(2001).
DR EMBL; AJ279031; CAC09416.1; -.
DR HSSP; P00750; 1TPG.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002258; NHL.
DR InterPro; IPR006530; YD.
DR Pfam; PF01436; NHL; 6.
DR Pfam; PF05593; RHL; repeat; 5.
DR SMART; SM00181; EGF; 6.
DR TIGRFAMs; TIGR01643; YD repeat_2x; 7.
DR PROSITE; PS00022; EGF_1; 8.
DR PROSITE; PS01186; EGF_2; 7.
KW EGF-like domain.
FT NON_TER 1
FT TER 2802
FT FT 2802
SQ SEQUENCE 2802 AA; 310745 MW; B1FBC2C84EDFA4B3 CRC64;

Query Match 64.6%; Score 9616.5; DB 13; Length 2802;
Best Local Similarity 61.3%; Pred.No. 0;
Matches 1775; Conservative 411; Mismatches 491; Indels 217; Gaps 27;

Qy 1 MDVKERYPSRLTERR-DABERYTSSADSEEGKAP-QKSYSSSETLKAYDQDARLAYGS 58
Db 1 MDIKDR-HRSLTRGRCGEKRYTSSLSLSDSDCRVPAQKSYSSSETLKAYCHDTRMYGN 59
Qy 59 RVKDIVPQAEFEFCRTGANFTLRELGLEEVTPPHGTLYRTDGLPQCGYSGAGSDADME 118
Db 60 RVSLVHRESDEFPQGTNF-LAELGICERS-PHRSYCGSDIGILHQGYSLTSGSDADS 118
Qy 119 ADTVLSEHPVLRGRSTRSCRSSCLSRANSNLTLTDTEHEN----- 161
Db 119 TEGGKSPHEAIRLWGRGIKSRSSGLSRNSALTTLTSDNENKSDDEDFHLSLSEKL 178
Qy 162 -----TETHD--- 166
Db 179 DRQTSWQOLAEKTNLSLRPIPTPTSSSSSLPSAQLPSSHNPPPVSCQMLDSDNTSHQIM 238

```

```

Qy 167 -----PGGLQNH-----ARLUTP-PPPLSHAHTPNQHAASI 197
Db 239 DTNPDEEFSPNSYLLRACSGPQQASSGSPSNHHSOSTLRPLPPPHNHSLS---HHHSSA 295
Qy 198 NSLNEGNTTPSN---PSPAPTOHSLSGEPAGGAQPAHAQENWLNLSNIPLETNLG 253
Db 296 NSLNSNSLNRNQIHAPAPND--LATP-----ESVOLQDSWLNLSNVPLETR--- 344
Qy 254 KOPFLGTLDNLIEMDILGASRHDGAYSDGHFLFK-PGQTSPLFCTTSPGYPLTSTSVYS 312
Db 345 -----HFLPKTSSGTTPLFSSSPGYPLTSGTVT 374
Qy 313 PPPRLPRSTPAPAPNLKPSKYCNWKCAALSAIVISATLVILLAYFVAMHFLGNWHL 372
Db 375 PPRLLPRNTFSRNAPLKLKPSKYCSWKCAALSAIAAAVLLAILLAYFIAMHLLGNWL 434
Qy 373 QMEQGMVEI-----TEDTASSVPVPTDVSYPSSGGTGLTDPDRKGKGTGKPSF 424
Db 435 QPADGHTFSNGLRPGCAAGAEDEGAARPPA-----GRG-----PW 467
Qy 425 FPEDFIDSGEIDVGRRASQKIPGTFWRSQVFDHPVHLKFNVSGLKAAALVIGYGRKL 484
Db 468 VTRNSSIDSGETEVRKVTQEVPGVFWRSQIHISQPQFLKFNISLGDALFGVYIRGL 527
Qy 485 PPSHTQEDFVELLDGRLLTOFARSLGTPRQSRGTVPSSSHETGFIQVLDSGIWHLAFY 544
Db 528 PPSHAQYDFMERLDGK-----EKWSVSPRERSIQTLVQNEAVFYVYLDVGLWHLAFY 582
Qy 545 NDGKSEVVSFLTTAIESVDNCPSCYNGNCISGTCCHFCFLGFLGPDGCRASCPCVLCGN 604
Db 583 NDGKKEVVVSTVILDSVQDCPRCHNGCEVSGVCHCFPGFHGADCAKAAACPVLCGN 642
Qy 605 GQYMKRCLCHSGWKGAEDUPTNOCIDVACSNHCTCTGTICNPGYKSGSCBVDQWD 664
Db 643 GQYSGTCLCYSGWKGPCDVPISQIDPSCGGHSGCIEGNCVCVSGIGKGENCEEVDCLD 702
Qy 665 PTCGRGVCVRGECHCFVGMGTNCETPRATCLDQCSGHGTPLPDTGLSCDPSFTGHDC 724
Db 703 PTCSNHGVCVNGECLCSFGWGGINCELPRAQCPDQCSGHGTPLSDTGLSCDPSNMWGPC 762
Qy 725 SIEICAADCGHGVCGTTCRCEGDGWMGAACDORACHPCAHEGTCRGGKCECSGWNGE 784
Db 763 SVEVCSVDCGTHGVGIGGACRCEGWTGVACDQVCHPRCTEHGTCCKGKCEGWNGE 822
Qy 785 HCTIAHYLDRVVKGPCGLCNGNRCCTLDLNGWHCVCOLGWRGAGCDTSMETACGDSKN 844
Db 823 HCTIGRQTTGTDTGCPDLGNGRCCTLGQNSQVCVCTGWRGPGCNVAMETSCADNKN 882
Qy 845 DGDGLVDCMDPCCLQPLCHINPLCLGSPNPLDIIQETQVPVVSQQLNHSFYDRIKFLVR 904
Db 883 EGDGLVDCLVPCCLQSTCQNSLLCRGSRDPLDIIQQSH--SGSPAVKSFYDRIKLLVGK 940
Qy 905 DSHIIPGENPFDGGHACVIRGOVMTSDGTPLAGVNI SFVNNPLFGYITSRDGSFDLVT 964
Db 941 DSHIIPGENPFSNLSLIRGQVVTDTGTPLVGNVSFVKYKPYGYITITRODGMFDLVA 1000
Qy 965 NGGISILIRFERAPITQHTLWLPWDRFFVVMETIMRHEENEIPSCDLSNFARPNPVVS 1024
Db 1001 NGGSLTLHFERAPFMSQERTVWLPWNSFYAMDTLVKMTENSIPSCDLSGVRPDPVII 1060
Qy 1025 PPLTSPASSCAEKPIPEIQAQOEIISICKMRSLSYLSRTPGYKSVLSLTHPTI 1084
Db 1061 SPSLTFPFDAPGRNPIVPTQVLHEEIEVPGSSIKLILYLSRTPGYKSVLSLTHPTI 1120
Qy 1085 PNLKMKVHLMVAVEGRLFKFWFAAADLSYVFIWKTQVYNQKVFGLSEAFVSQVEYES 1144
Db 1121 PNLKLVHLMVAVEGHLFKSFLASNLAYTIFWKTDAYGQVYGLSDAVVSQVEYET 1180
Qy 1145 CPDLILWEKRTTVLQGYEIDASKLGWSLDKHALNIOGILHKGNGENQFVSQPPVIG 1204
Db 1181 CPSLILWEKRTALQGFELDPNLSGWSLDKHHVNLVKSGLHKGNGENQFVTOQPAVIT 1240

```

```
QY 1205 SIMNGRRRSISCPSCNGLADGNKLLAPVALTCSDGLXVGFNFYIRIRIPPSGNVTNLL 1264
Db 1241 SIMNGRRRSISCPSCNGLAEGNKLAPVALAVGIDGLFVGDFNITRIRIPPSGNVTNLL 1300
QY 1265 ELRNKDRFHSHPAHKYLLATDPMGAVFVSDNSRREVKIKSTVVVVKOLAVKNSVAVGT 1324
Db 1301 ELRNKDFHSHNPAHKYLLAVDPVSGSLYSDTNSRIYKVSLSLTGKDLAGNSEVAVGT 1360
QY 1325 GQCLPFDDTCRGDGGKATEATLNPRTGIVDFGLTYFDGTMIRRIDQNGIISTLLGS 1384
Db 1361 GEQCLPFDEARCGDGGKAVDATLMSPGIAVDKYLGMVFDATMIRKVDQNGIISTLLGS 1420
QY 1385 NDLTSARPLSCDSVMDISQVRLEHPTDLAINPMDNSLYVLDNNVQLQISENHQVRIVAGR 1444
Db 1421 NDLTAVRPLSCDSMDVSVQVLEHPTDLAVPMDNSLYVLENNVILRITENHQVSIAGR 1480
QY 1445 PMHCQVPGIDHFLSKVAIHATIESATALAVSHNGVLYIAETDEKKNIRIQVTTSGEIS 1504
Db 1481 PMHCQVPGID-YSLSKLAIHSALESASAIASHTGVLVYISETDEKKNIRLQVTTNGEIC 1539
QY 1505 LVAGAPGCDKNDANCDSCFDDGYAKADKLNTPSSSLAVCAVAGELYVADLGNIRFRIR 1564
Db 1540 LLGAASDCDCKNDVNCYSGDDGYATDAILNSPSSSLAVAPDGTIYIADLGNIRIRAVS 1599
QY 1565 KNRKPLANTONMYELSSPIDELYLFDTTGKHLVYTSQSLPTGDLVYNTFTYTGDDITLITN 1624
Db 1600 KNRFILMSFQYEAASPGEGELYVFNADGHHQVTLVLTGELYLYNTYSSDNDVTEVMS 1659
QY 1625 NGNMVYVRDSTGMPLVLVDPGVQVYVMTGNSALKSVTTQGHELAMTYHNGSGLLAT 1684
Db 1660 NGNSLKVRDASGMPRHLLPNDQIVTLAVGTNGGLKLVSTQTLGLTMYNGSGLLAT 1719
QY 1685 KSNENCHTTFEYDVSFGRILNVTFTPTQVSVSPSDTSSVHVQVETSSK-DVVTITNLS 1743
Db 1720 KSDETGWTTFYDHEGRLLNVTFTPTQVSVSLHREMEKSTIDIENSRRDDVDVVTINLS 1779
QY 1744 ASGAFYTLLOQVNRNSYIIGADGSLRLLLANGNEVALQTPHLLAGTVNPTVGRKRVNTP 1803
Db 1780 SVEASYTVQDQVNRNSYQLCNGNGLTRVMYANGNSISFHPSEPHVLAGTVPTIGRCNISLP 1839
QY 1804 IDGNLNVHRQKQEQVTVFGRRLRVHNRNLLSLDPRVTRTEKYIDHHRKFTLRI 1863
Db 1840 MENGLSIEVRLKEQIKGKTVFGRKLRVHGNLLSIDYDRNIRTEKYIDHHRKFTLRI 1899
QY 1864 LYDQAGPSSLWSPSSRLVNVVTVSPGYIAGIQGIMSERMEYDOAGRITSIPADGKT 1923
Db 1900 IYDQGRPFLWLPSSGLAANVSYFFNGRLAGLQRGAMSERTDIDQGRILISEMFPADGKV 1959
QY 1924 WSYTYLEKSMVLLHLSQRYIFBFDKNDRLSVVTPMNVARQTLTIRSVGYRNYIYQPP 1983
Db 1960 WSYTYLEKSMVLLHLSQRYIFBFDKNDRLSVVTPMNVARQTLTIRSVGYRNYIYQPP 2019
QY 1984 GNASVIOFTEDGHLLATFVLGTRRVYVYKGLSKLAETLYDTTKVSYETDETACMLKT 2043
Db 2020 SNASVIFDSDGRILTSFLGTRQRYFYKGLSKLSEIVYSTAVTGYDETTCVGLKN 2079
QY 2044 INLQNEGTCTIRYQIGPLIDIQIFRFTPEGMVNFARFDYNY-DNSFRVTSQVAVINETP 2102
Db 2080 VNLQSGGFSCTIRYKIGPLVDQIYRFSSEGMVNFARFDYVHDNSFRIASIKPIISETP 2139
QY 2103 LPDLVYDDVSGKTEQFGFVGYIYDINOIITAVMTHTKHDAVGRMKEVQYELFRSL 2162
Db 2140 LPVDLYRDEISGKVEHFGFVGYIYDINOIITAVMTLSKHDFTHGRKEVQYEMFRSL 2199
QY 2163 MYMTVQYDNNGRVVKELKVGYPANTRYSYEYDADGQLQTVSINDKPLMRYSYDLNCG 2222
Db 2200 MYMTVQYDNNGRVVKELKVGYPANTRYSYEYDADGQLQTVSINDKPLMRYSYDLNCG 2259
QY 2223 LHLLSPNSARLPLRVLDIDRTRIGDVOYKMDGDFLRQRCGDIFFENYSGALLIKAYN 2282
Db 2260 LHLLPNPNSVRLPLRVLDIDRTRIGDIPYKIDDDGFLCQSGDSVFEYNSKGLLATRAYN 2319
QY 2283 RAGSWSVRYRVDGLGRVRSKSSHLLQPFYADLTNPTKVTHLYNHSSEITSLSYDLDQ 2342
```

```
Db 2320 KANGNVQYRYDGLGRASCNTLGHLELQYFADLHNPTRVTHVYNSHSEITSLSYDLDQ 2379
QY 2343 GHLFAMELSSGDEFYIACDNTGTPLAVFSGTGLMIKOILYTAYGEIYMDTNPFIILIGY 2402
Db 2380 GHLFAMESSSGSEYVAVSDNTGTPLAVFSINGLMIKOLQYTAIGEYIYDNPFIILIGY 2439
QY 2403 HGGYDPLTKLVHMRDYYDLVAGRTSPDHELKHLSSSNVMPFNLYMFRNNPINSQ 2462
Db 2440 HGGYDPLTKLVHMRDYYDLVAGRTSPDHELKHLSSSNVMPFNLYMFRNNPINSQ 2498
QY 2463 DIKCFMTDVSMLLTFGQLHNVIPGYKPKDMDAWEPSYELIHTQMKTOEMDASKSLGV 2522
Db 2499 DLKNYVTDVKSMLVWFGQLSNIIPGFRKKNYFVSPPEL----TESQACENGQLITGV 2554
QY 2523 QCEVQKQLKAFVTLERFDQLYSGTITSCQAPKTK----FASGSGVFGVGVFALKDGR 2578
Db 2555 QOTTERHQAFALE-----GVISKRLHASIREKAGHWFATSTPIIGKIMFAVKGR 2608
QY 2579 VTTDIISVANEDGRVAAILNHAHYLENLHPTIGDVTHTYFVKPSEGDILAILGLSGR 2638
Db 2609 VTTGISSATDSDRKIASVLNSAHYLEKMHYSIEGKDTYFVKIGSADSLVTLAMTSGR 2668
QY 2639 RTLENGVNVTVSQINTVLNGRTRYTDIQLQYCALCLNTRYG---TTLDSEKARVLSLAR 2695
Db 2669 KVLDSGVNVTVSQPTLLINGTRFRFTNIEFQYSTLLNIRYGLTADTLDSEKARVLDQAR 2728
QY 2696 QRAVQAWARQOBLRGEGLRAWTEGEGKQOVLSTORVQYDGFVSVIEQIPELSDSA 2755
Db 2729 QRALGSAAWAKQOQKARDGREGSVMTDGEKQQLNTGRVQYEGYVLPVEQIPELADSS 2788
QY 2756 NNHFMROSENGRR 2769
Db 2789 SNIQFLRQNGMKR 2802

RESULT 8
Q9WTS5 PRELIMINARY; PRT; 2764 AA.
AC Q9WTS5,
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Ten-m2.
GN OD22 OR TEN-M2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c; TISSUE=Brain;
RA Coshishi T., Zhou X., Feng K., Richter B., Moergelin M., Perez M.T.,
RA Su W., Chiquet-Ehrismann R., Rauch U., Faessler R.;
RT "Mouse Ten-m/Odz is a new family of dimeric type II transmembrane
RT proteins expressed in many tissues.";
RL J. Cell Biol. 0:0-0(1999).
DR EMBL; AB025411; BAA77397.1; -.
DR HSSP; P35555; 1EMN.
DR MGD; MGI:1345184; Odz2.
DR GO; GO:0005198; F:Structural molecule activity; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR006530; YD.
DR Pfam; PF00008; EGF; 4.
DR Pfam; PF01436; NHL; 6.
DR Pfam; PF05593; RHL-repeat; 6.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00181; EGF; 5.
DR TIGRFAMs; TIGR01643; YD repeat_2x; 7.
DR PROSITE; PS00022; EGF_1; 8.
```

DR PROSITE; PS01186; EGF_2; 7.
KW EGF-like domain.
SQ SEQUENCE 2764 AA; 306464 MW; 73BA3D916D0F0344 CRC64;

Query Match
Best local similarity 64.6%; Score 9610.5; DB 11; Length 2764;
Matches 1769; Conservative 418; Mismatches 489; Indels 181; Gaps 28;

QY 1 MDVKRKPYRLRR-DAERRYTSSADSEEGAP-QKSYSSSETLKAYDODARLAYGS 58
DB 1 MDVKRR-HRLTRGCKEYVTSLSLSEDCRVTQKSYSSSETLKAYDSDRMHYN 59

QY 59 RVKDIYQRAEPCRTGANTFRELGLEVPHPHGLYTDGLPCGYSMGAGSDAME 118
DB 60 RVTDLHRESDEFSROCTNFTLAELGICEPS-PHRSYCSMDGILHQYSLSGTSDSD 118

QY 119 ADVLSPHPVRLWLGSTSGRSSCLSSRANSNLTDTEHENTEDHPG----- 168
DB 119 TEGGMPFAIRLWGRGIKSRSSGLSRNSALTUTSDNENKSDDDNGRPIPTSSSS 178

QY 169 -----G 169
DB 179 LLPSAQLPSSHPVSCQPLDLSNTSHQIMDNPDEEFPSPNSYLLRACSGPQQAASSG 238

QY 170 LQNH---ALRTTTPPLSHAHTNOHHAASINSLEGNTPRSN-----PSPAPTHSLSG 222
DB 239 PPHHSQSTLRPLPP-PHNHTLUSHH-SSANSLENSLTNRSSQIHAPAPND--LAT 294

QY 223 EPPAGGAQEPAAQENWLLNSNIPLETRNLGKOPFLGLQDNLIEMDILGASRHDGAYSD 282
DB 295 TP-----ESVQLQDSWLSNVPLETR----- 316

QY 283 GHFLFK-PGTSPLFCTTSPGYPLTSTVYSPPPRPLRSTPARPAPNFKPKSKYCNWKC 341
DB 317 -HFLFTSSGSTPLFSSSPGYPLTSGTVTPRLLPRNTFSKAFKLKPKSKYCSWK 375

QY 342 AALSAIVISATLVIILAYFVAMHFLGNHQLQWEGMEYIETEDTASSWPEVPTDSLPS 401
DB 376 AALSAIAALLIALLIAYFIAMLLGNLQWQADGHTF--NGVVTGLPGNDDVATPS 433

QY 402 GGTGLETPRDKGTTGTEKPSPPEDSFIDSGEIDVGRRAQKIPPGTFWRSQVFIHP 461
DB 434 GGR-----VPSL--KNSSIDSGEAEVGRVTVQEVPPGVFWRSQIHISQ 476

QY 462 VHLKFNVLGKAALVGIYGRKGHPSSHOTQDFVELLDGRLLTQEARSLGCTPROSGTV 521
DB 477 QFLKFNISLGKDALFGVYIRRGUPPSHAQYDFMERLDGX-----EKWSVESPRRRSIQ 531

QY 522 PPSSETGTQYLDSGIWHIAFYNDGKESEWVSFLATAIESVDNCPSCNCGNDCISGTC 581
DB 532 TLVQNEAVFYQLDVLGLHLAFYNDGKDEWVSFNTVVLDSVDQCPNCHGNGEVCGLC 591

QY 582 HCFGLFLGPDGCRASCPLVCSGNQWNGRCLCHSKWGAECDVPTNQCIDVACSNHGT 641
DB 592 HCFPFLGADCAKACPLVCSGNQYSGKTCQCSGWKGAECVPMNQCIDPSCGHGSC 651

QY 642 ITGTCICNPYKGESEEDVMDPTCSGRGVVGRGCHCFVGMGNGTCTPRATCLDQCS 701
DB 652 IDGNCVCAAGYKGEHEEVDCLDPTCSSHGVGVNGECLCSPGNGGLNCLARVQCPDQCS 711

QY 702 GHGTFPLDTCSCDPSWTGHDCSIICADCGHGVGVGCTCRCDGWMGAACDORACH 761
DB 712 GHGTYLPDSGLCSDPNWMPDCSV--VCSVDCGTHGVCIGGACRCBEGTGAACDQVCH 770

QY 762 PRCAEHTCDGKCECSPGWNGEHCITIAHYLDRAVKEGCGGLCNGNGRCTLIDNGHVCV 821
DB 771 PRCEHGTCKDCKECCRGWNGEHCIT-----DGCPLDLCNGNGRCTLGQNSWQVCV 821

QY 822 QLGWRGAGCDSMETACGSKNDGDLVDCMDPDCCLOPLCHINPLCLGSPNPLDIIQE 881
DB 822 QTGWRGPGCNVAMETSCADKNDEGDLVDCLDPDCCLOSAQCNLSLCRGRDPLDIIQ 881

QY 882 TQVPVSOQNLHSFYDRIKFLVGRDSTHIIIPGENPFDGGHACVIRGQWTSDDGTPLVGNI 941

DB 882 GQ--TDWPAVKSFYDRIKLAGKDSHIIIPGDNFPNSSLVSLRQVQVMTDGTPLVGNI 939

QY 942 SFVNPPLFGYTIISRODGSFLVTNGGISILRFRAPFIQEHTLWLPWDEFFWETILM 1001

DB 940 SFVKPKGYTITIQDGTFLDIANGSALTLHFRAPFMQSERIVLWLPWNSFYANDILVM 999

QY 1002 RHENEIPSCDLSNFARPNPVWSPSLTSPASSCAEKGPIVPEIQALCEEISISGCMRL 1061

DB 1000 KTEENSIPSCDLSGFVRPDPPIIISPLSTFSPASPNPIVPEIQVLEHIELPQTNVKL 1059

QY 1062 SYLSRTPGYKSVLRISLTHPTIPFNLMKHVLMAVEGRLPKFWFAAAPDLSYFINDKT 1121

DB 1060 RYLSRTAGYKSLKITITQSTVFLNLTIRVHMAVEGHLFQKSFQASPNLAYFINDKT 1119

QY 1122 DVYNQKVPGLSEAFVSVGYEYESCPDLILMEKRTTVLQGYEIDASKLGGSLDXHALNI 1181

DB 1120 DAYQRYVGLSDAVVSVGFVETCPSLILMEKRTALLQGFELDSPNLGGSLDXHTLNV 1179

QY 1182 QSGILHKGNGNQVSOQPPVIGSIMNGRRRSISCPSCNGLADGNKLLAPVALUTCSDG 1241

DB 1180 KSGILHKGNGENQFLTOQPAIITSIMNGRRRSISCPSCNGLAEGNKLLAPVALAVIDG 1239

QY 1242 SLVYGDENYIRRIIPSGNVTNILELRNKDPFHSHSPAKYLYLATDPMGSAVFLSDSNRR 1301

DB 1240 SLFVGDFNYIRRIIPSRNVTISILELRNKEFKHNSPGHKYLYLAVDPVTGSLYVSDTNSRR 1299

QY 1302 VFKIKTVVVXKDLVKNSEWAGTGDQCLPFDDTRCGDGGKATEATLTNPRGITVDKFLGI 1361

DB 1300 IYRVKLSGAKDLAGNSEWAGTGEQCLPFDEARCGDGGKAVDATLMSPRGIADVKNGLM 1359

QY 1362 YFVGTMTIRRDONGIISTLLGSNDLTSARPLSCDSVMDISQVLEPTDLAINPMNSL 1421

DB 1360 YFVDATMIRKYDONGIISTLLGSNDLTAVRPLSCDSVDVAQVLEWPTDLAVPMNSL 1419

QY 1422 YVLDNNVLOIASENHQVRIVAGRPMHCOVPGCIDHFLLSKVAIHATLSATALAVSHNGVL 1481

DB 1420 YVLENNVLRITENHQVSIAGRPMHCOVPGCID-YLSLKLAHSALESASAIASHTGVL 1478

QY 1482 YIAETDEKKNIRIQVTTSGEISIVAGAPSCDCKNDANCDPFGDDGYAKADKLNPSS 1541

DB 1479 YITETDEKKNIRLRQVTTNGEICILLAGAASDCDCKNDVNCICYSGDDAYATDAILNSPSS 1538

QY 1542 LAVCADGLYVADLGNIRIRPIRNKPLNTQNMVELSSPIDQBLXLYLFTDTGKHLTYOSL 1601

DB 1539 LAVAPDGTIYADLGNIRIRAVSKNPKVLNAFNQYEAASFGQEQLYVFNADGIHQYTVSL 1598

QY 1602 PTGDYLYNFTYTGDCDITLITDNNGMVNVRRDSTGMPLMLVVPDGGVYVMTGINSALK 1661

DB 1599 VTGEYLYNFTYSADNDVTLELDNNGSLKIRRDSSGMPRHLLMPDNQIITLTGVTNGGLK 1658

QY 1662 SVTTQGHLEAMWYHNSGLLTKSNENGTTFEYDSFGRLTNVTPTGOVSSFRSDTD 1721

DB 1659 AVSTQNUELGMTYDGNITGLLATKSDGTGTFYDYDHEGRLTNVTPTGOVSSFRSDTD 1718

QY 1722 SSVHVQVETSSK-DDVTITNLSASGAPYTLLOQVNSYVIGADGSLRLLLANGMEVAL 1780

DB 1719 KSIITDIENSNRDDVTITNLSVEASYSYVQVENSQYLCNNGTLRVMYANGMAVSF 1778

QY 1781 QTEPHLAGTVNPTVGRKNVTLPIDNGLNVVWQORKEAQGVTVVGRRLRVNHNLLS 1840

DB 1779 HSEPHVLAGTITPTIGRCNISLPMENGLNSIEWLRKEQIKGKVTIFGRKLVRHGRNLS 1838

QY 1841 LDFRVTETEKIYDHRKFTLRILYDQAGRPSLSPSSRLNGVNVTVSPGGYIAGIQSGI 1900

DB 1839 IDYDRNIRTEKIDYHRKFTLRILYDQVGRFPLWLPSSGLAANVSVFFNGRLAQORGA 1898

QY 1901 MSERMEYDQAGRITSRIFADGKTWSTYLYEKSVMLLLSHQRYIPEFDKNDRLSSVTNPN 1960

DB 1899 MSERTDIDKQRIVSERPMFADGKWSYSLDKSVLLLSQSRQYIFEYDSDSRLEHVTMPS 1958

QY 1961 VARQLETIRSVGYRNIIYQPEGNASVIQDFTEDGHLHTFTYLGTERVLYKYGKLSKL 2020

Db 1959 VARHSMSTHTSICGIRNINYPNESNASVIFDYSDDGRILKTSFLGTGRGVFYKYGKLSKL 2018
Qy 2021 AETLYDTTKVSFYDETAGMLKININQNGSFTCTIRYQIGELIDROIIFRPFEBGMWNR 2080
Db 2019 SEIYVSTAVTFQYDETTGLKWNVIQSGSFTCTIRYKVGVLVDKQIYRFESEGMNAR 2078
Qy 2081 FDYNY-DNSPRVTSMAVINETPLPDLVRYDDVSKTEQFGKFGYIYDINQIITAM 2139
Db 2079 FDYTHDNSPRIASIKPVISETPLVDLYRDEISKVEHFKFGYIYDINQIITAM 2138
Qy 2140 THTKHFDAYGRMKEVQYEIFRSLMYMTVOYDMGRVVKELKVGPIYANTTYSYEYDAD 2199
Db 2139 TLSKHFDTHGRKEVQYENFRSLMYMTVOYDSMGRVVKRELKLGPIYANTTYYTYDYDGD 2198
Qy 2200 GOLQTVSINDKPLWRYSYDLNGNLHLLSPCNSARLTPLAYDIRDRITRLGDVQYKMDG 2259
Db 2199 GQJQSVAVNDRPTWRYSYDLNGNLHLLSPCNSARLTPLAYDIRDRITRLGDVQYKIDDDG 2258
Qy 2260 FLQRGSDIFEYNSAGLLIKAYNRAGSWRYRYDGLGRVRSKSSHHLQFFYADLTN 2319
Db 2259 YLCQSGSDIFEYNSKGLTRAYNKASGWSQYRYDGVGRASVYKTNLGHLLQYFYSDLHN 2318
Qy 2320 PTKVTHLYNHSSEITSLYDLOGLHFAVELSGDEFYIACDNIGTFLAVFGTGLMIKQ 2379
Db 2319 PTRITHVYHNSSEITSLYDLOGLHFAVELSGDEFYIACDNIGTFLAVFGTGLMIKQ 2378
Qy 2380 ILYTAYGEIYMDTNPNFQIIIGYHGLYDPLTKLVHMGRRDYDVLGRWTSDDHMLKHL 2439
Db 2379 LQYTAYGEIYDNPDPQWVIGFHGGLYDPLTKLVHFTQDQYDVLGRWTSFDTYWNV 2438
Qy 2440 SSSNVNPNLYMPKRNPNINSODIKCFMTDVNSWLLTFQFQHLNVIPQYKPDNDAMEP 2499
Db 2439 GKPEA-PFNLYMPKRNPNINSDELDKNYVTVDKSWLWVMEFQLSNIIPGPRAKYFVFP 2497
Qy 2500 SYELIHTOMKTQBWMDKSKSLTGVCQVQKQAFVTLERFDQLYGSTITSQQCAQPKTK- 2558
Db 2498 PYELSESQAS-----ENGQITGVQOQTERHNOQFALE-----GOVITKLUHASIREKA 2547
Qy 2559 ---FASGSGVFGKVPALKDGRVTTDIIISVANEDGRRVAAILHAHLYENLHFTIDGVD 2615
Db 2548 GHWFAITPIIGKGINFAIKEGRVTTGVSSIASERSKVASVNNAYLKDWHYSIEGKD 2607
Qy 2616 THYFKVSGSEGLALILGSGGRTLENGVNTVTSQINTVLNGRTRYTDIOLQALCL 2675
Db 2608 THYFVXIGAADGLVLTGTIGRKVLSEGVNTVTSQITLLVNGRTRYTNISFYSTLL 2667
Qy 2676 NTRYG---TTLDEKARVLELQARVQAWAREQORLEGEGLRAWTEGEKQOVLSTG 2732
Db 2668 SIRYGLTPDLDDEKARVLDQAGRALGTAWAKEQKQKARDGREGSELWTEGEKQQLSTG 2727
Qy 2733 RVQGYDGFVIVEQYPELSDSNHNFHRSQEMGR 2769
Db 2728 RVQGYGYVLPVEQYPELADSSNIQFLRQEMGR 2764

RESULT 9

Q9WTS4 PRELIMINARY; PAT; 2731 AA.
AC Q9WTS4;
DT 01-NOV-1999 (T-EMBLrel. 12, Created)
DT 01-NOV-1999 (T-EMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Ten-m1.
GN CDZ1 OR TEN-M1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Balb/c;
RA Ohashi T., Zhou X., Feng K., Richter B., Moergelin M., Perez M.T.,
RA Su W., Chiquet-Ehrismann R., Rauch U., Faessler R.;

"Mouse Ten-m/Odz is a new family of dimeric type II transmembrane proteins expressed in many tissues.";
RL J. Cell Biol. 0:0-0(1999).
EMBL; AB025410; BAA77396.1; -.
HSSP; P00750; 1TFP.
DR MGD; MGI:1345185; Odz1.
GO; GO:0005887; C:integral to plasma membrane; IDA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; EGF.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR006530; YD.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF01436; NHL; 5.
DR Pfam; PF05593; RRS repeat; 6.
DR SMART; SM00181; EGF; 7.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 7.
DR PROSITE; PS00022; EGF_1; 8.
DR PROSITE; PS01186; EGF_2; 7.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
KW EGF-like domain.
SQ SEQUENCE 2731 AA; 305752 MW; 9129PA4CFE4A7770 CRC64;
Query Match 64.5%; Score 9607; DB 11; Length 2731;
Best Local Similarity 62.5%; Pred.No.0;
Matches 1754; Conservative 407; Mismatches 533; Indels 112; Gaps 19;
Qy 1 MDVKERKPYRSLTR-RSDAERRVYTSSSADSEEGKAPKSYSSSETLKAYDQDARLAYGS- 58
Db 1 MEQTDCEPYQPLSKVKHEMDLAYTSSDESEDEGRKQSFNSRETLHYNQELRRYNSQ 60
Qy 59 --RVKQIVQEAQ-EFCRTGANFTLRELGLVEVTPPHGTL---YRTDI-GLPCQGYSMGA 111
Db 61 SRKRDVKEKSTQIEFCE-----TTP--TLCSGYHTDMESVSRHGQLEM 103
Qy 112 GSDADMEADTVLSPHVPVRLGWSRSGSSCLSSRANSNLTLTDTEHENTETDHPGGLQ 171
Db 104 GSDVDTTEGAASPDHALRWIRGMKSEHSSCLSSRANSALSUTDIDHER-KSDGNGFK 162
Qy 172 -----NHARLR--TPPPLSHAHTENQHAASINSLNRGNT 206
Db 163 FSPVCCDMEAPADSADQMSSPHNQFTFRPLPPPPHACTCARPPPTVDSLQRRSMT 222
Qy 207 PRSNPSAPPTDLSLSEPPAGGAQEPAAQENWLLNSNPLETNLNGKQPLGLTLDNLI 266
Db 223 TRSQFSPA-----APAPPTSTQDSVHLHNSWVLSNIPLETR----- 259
Qy 267 EMDILGASRDGAYSDGHFLFKPG-GTSPLCFTTSQYPLTSTVYSPPPRPLPRSTFAR 325
Db 260 -----HFLFKHSGSSAIFSAASQNYPLTSNTVYSPPPRPLPRSTFAR 302
Qy 326 PAFNLKPSKYCNWKAALSIAIVISATLVLLAYFVAMHLFGLNHLQPMEGOMYITED 385
Db 303 PAFTFNKPYRCCNMKCTALSATAITVTALLLAYVIAVHLFGLTQLPV-GQIYANGIS 361
Qy 386 TASWPVPTDVSLYPSGGTGLTEPDRKGTTEGKSPSPPEPDSFIDSGEIDVGRASOK 445
Db 362 NGNPCTESMDTTSYPIGGRVSDKSEK-----VFQKGRADTDEVDIGAQMVT 410
Qy 446 IPPGTFWRSQVFIIDHPVHLKFNVSGLKAAALVGIVYGRKGLPPSHPTQDFVELLDGRLLTQ 505
Db 411 IPPGLFWRFQITTHHPYILKFNISLAKDSLLGIYGRNIPPTHTQDFVKMLDGKQLVKQ 470
Qy 506 EARSLEGTQPSRGTVPPSSHETGFTOYLDSDGWHLAFYNDGKESVSVFLTAIESVDN 565
Db 471 DKSESDDIQHSRPNLILTSLETQFTIEYMDQGWLYAFYNDGKMEQVFLITAIEMDD 530
Qy 566 CPNVCYNGDCISGTCHCFLGFLGPCGRASCPCVLCSGNGQYMKGRCLCHSGWKGAECDD 625
Db 531 CSTNCNGGECISGHCHCFFPGFLGPCARDSCPCVLCGGNGEYKGHGKVCVRNGWKGPEDCV 590
Qy 626 PTQCIDVACSNHGTCTGTCTICNPGYKGESCEVDCMDPTCSGRGVCVVRGECHCFVNG 685

Db 591 PBEQIDPTCFHGHTCIMGVCI VPGYKGEI CBEEDCLDPMCSSHGI CVKGBCHCSTGWG 650
 Qy 686 GTNCEPPTATCLDOCSGHGTFLDPDTGLSCDPSWTGHDSCIEI CAADCGHGVGVGTOR 745
 Db 651 GVNCEPPLPTCQBQCSGHGTFLDGTGVCSDDPFWTGTSCSTELCTMECCSHGVCSSRGICQ 710
 Qy 746 CEDGWMGAACDORACHPRCAEHGTCDRGKCECSPGWNGEHTTIAHYLDLRVVKGCPCGLNC 805
 Db 711 CEEGWGPTCEERSCHSHCAEHGCKDGKCECSPGWNGEHTTIAHYLD -AVRDGCEGLCF 769
 Qy 806 GNGRCTLDLNGHVCQOLGWRGAGCDTSMETACGDSKNDNDGDLVDCMPDDCCLQPLCHI 865
 Db 770 GNGRCTLDQNGHVCQOVGSGTGCNIVMEMLCGNDLNDGDLTDCVDPDCQCCQNCYV 829
 Qy 866 NPLCLGSPNPLDIIQTOVPSQNLHLSFYDRIKELVGRDSTHIIPEGNPFDCGHACVIR 925
 Db 830 SPJCQSGPDDLIQSQPFLFSQHTSEFLFYDRIKELVGRDSTHIIPEGNPFDCGHACVIR 889
 Qy 926 GOVMTSDGTPLGVNLSFVNNPLFGYTIISRDQSGFDLVNTGGSIIILREFRAPFIQERT 985
 Db 890 GOVAVDGTPLGVNLSFVNNPLFGYTIISRDQSGFDLVNTGGSIIILREFRAPFIQERT 949
 Qy 986 LWLPDWRFFVMEIIMRHENEIIPSCDLSNFARPFPVPSPLTSFASSCAEKGPVPEI 1045
 Db 950 LWLPWNOFIVVEKVMQRIWADAPSCDISNFISPNFIVPSPLTSFASSCAEKGPVPEI 1009
 Qy 1046 QALQEEISIGCKMRLSYLSRTPGYKSVLRISLTHPTIPFNLMKVLHMAVEGRFRKW 1105
 Db 1010 QVQBEIPISSFVRLSYLSRTPGYKSVLRISLTHPTIPFNLMKVLHMAVEGRFRKW 1069
 Qy 1106 FAAAADLSYFVNDKTVNOKVFGJSEAFVSVGYEYESCPDILMEKRTVLQGYEIDA 1165
 Db 1070 FPAAINLVYFAWNTKDIYQKVWGLAEALVSVGYEYESCPDILMEKRTVLQGYEIDA 1129
 Qy 1166 SKLGGWLSKXHALNIOGILHKGNGENQVSOQPPVIGSIMGRRRSTSCPSONGLAD 1225
 Db 1130 SNLGGWLSKXHALNIOGILHKGNGENQVSOQPPVIGSIMGRRRSTSCPSONGLAD 1189
 Qy 1226 GNKLAPVALTCSGSLYVGFENYIRIPPSGNVTNILELRNKPFRHSHSPAKHYLAT 1285
 Db 1190 NNKLAPVALTCSGSLYVGFENYIRIPPSGNVTNILELRNKPFRHSHSPAKHYLAT 1249
 Qy 1286 DPMGAVFLDSNSRRVFKISTVVVDLVNKEVAVAGTDCQCLPDDTCGCGKATEA 1345
 Db 1250 DPMSESLYSDTNRKYVYKLSLVETKXLSNFEVAVAGTDCQCLPDDTCGCGKATEA 1309
 Qy 1346 TLTPRGITVDKFLIYFVDTMIRRIDONGIITSLGSDNLTSAKPLSCDSVMDISQVR 1405
 Db 1310 SLNSPRGITVDRHGFIVFVDTMIRRIDENAVITVIGSNGLTSTQPLSCDSVMDISQVR 1369
 Qy 1406 LEWPTDLAIPMDNSLVLDNNVLOISENHVRIVAGRMHCOVPGIDHFLSKVAIRA 1465
 Db 1370 LEWPTDLAIPMDNSLVLDNNVLOISENHVRIVAGRMHCOVPGIDHFLSKVAIRA 1429
 Qy 1466 TLESATALAYSHNGVLIAETDEKINRIQVTTSGTSGEISLVAGAPSCDCKNDANDCF 1525
 Db 1430 TLESARAISSHGLLIATDEKINRIQVTTSGTSGEISLVAGAPSCDCKNDANDCF 1489
 Qy 1526 GDDYAKDANKLTPSSLAVCADGELYVADLGNIRIRIRKPKPLANTQNMVYESSPIDQ 1585
 Db 1490 GDGCVAKDANKLTPSSLAVCADGELYVADLGNIRIRIRKPKPLANTQNMVYESSPIDQ 1549
 Qy 1586 LYLFDTTGKHYTOSLITGTYLYNFTTGGDDITLITDNGNVMVVRDSTGMLVLP 1645
 Db 1550 LYQFTVNGTHLHTNLTTRDYVYNTYNAEGDLGALITSSNGNSVHIRRDAGGMLVLP 1609
 Qy 1646 DGQVYVMTGINSKLVTTQGHBLAMWTHYHNSGLIATKSNENGTTFEYEDSFGRLTN 1705
 Db 1610 GGQVYVMTGINSKLVTTQGHBLAMWTHYHNSGLIATKSNENGTTFEYEDSFGRLTN 1669
 Qy 1706 VTPPTQGVSSFRSDTSVHVQVETSKDDVTTITNLSAGAFVTLTQDVRNSYIGAD 1765
 Db 1670 ATFPPTGEVSSHSDLEKLTVALDTSNRENVLSTNLATSTIYILKQENTQSTYRVS 1729

RESULT 10
 Q9RIK2
 ID Q9RIK2 PRELIMINARY; PR: 2765 AA.

Qy 1766 GSRLLLLANGMEVALQTEPHLLAGTAVNPTVCKRNVTLPIDNGLNLVEMRQKEARGOVT 1825
 Db 1730 GSRVTFASGMEINLSSEPHLAGAVNPTLCKGNISLPGHEANLIEWRQKEONKGVNS 1789
 Qy 1826 VFGRLRLVHNENLALSDPDRVTRTEKIYDDHRKFTLRLYDQAGRPSLWSPSSRLNGVY 1885
 Db 1790 AFERELRAHNENLALSDPDRVTRTEKIYDDHRKFTLRLYDQAGRPSLWSPSSRLNGVY 1849
 Qy 1886 TYPGQYTAGTORGIMSERMEYDQAGRTSIFADGKTWSYILEKSVILLHSHORQVIF 1945
 Db 1850 TYPSEGLVTFIQRTWNEKMEYDQSGKIIISRTWADGKTWSYILEKSVILLHSHORQVIF 1909
 Qy 1946 BFDKNDRSSVTMPNVAROTLETIRSVGYRNIYQPPPEGNASVIOQFTEDEGHLHFTYLG 2005
 Db 1910 EYDQSDCLLVTMESVHSLQTLMSVGYRNIYTPPDSSTSFQDYSDRDEGLLQTLHLG 1969
 Qy 2006 TORRVIYKXKSLAETLYDTTKVSTFYDTAGMKTINLONSGFTCTIYRQIGPLID 2065
 Db 1970 TORRVIYKXKSLAETLYDTTKVSTFYDTAGMKTINLONSGFTCTIYRQIGPLID 2029
 Qy 2066 ROIFPTEEGVNARFDYNDSPRVTSMQAVINETPLPIDLYRVDVSGKTEQFGKGV 2125
 Db 2030 ROIFPTEEGVNARFDYNDSPRVTSMQAVINETPLPIDLYRVDVSGKTEQFGKGV 2088
 Qy 2126 IYDINQIITAVMTHTKHFDAYGRMKEVQYEIFSLMYMVTQYDNNMGRVVKELKVP 2185
 Db 2089 INYDINQIITAVMTHTKHFDAYGRMKEVQYEIFSLMYMVTQYDNNMGRVVKELKVP 2148
 Qy 2186 YANTTRYEYVDADQLOTVSINDKPLWRYSDVLDNGLHLLSPGNSARLTPLRYDIDRI 2245
 Db 2149 DANITRYEYVDADQLOTVSINDKPLWRYSDVLDNGLHLLSPGNSARLTPLRYDIDRI 2208
 Qy 2246 FRLGVOYKMBDEGFLRQGGDIFEYNSAGLLIKAYNRAGSVRYRYDGLRRVSSKSS 2305
 Db 2209 FRLGVOYKMBDEGFLRQGGDIFEYNSAGLLIKAYNRAGSVRYRYDGLRRVSSKSS 2268
 Qy 2306 HSHHQFYADLTNTKVTHTLYNHSSEITSLYDLOGLHFAELSSGDEYFIACDNTGT 2365
 Db 2269 LQCHLQFFYADLTNTKVTHTLYNHSSEITSLYDLOGLHFAELSSGDEYFIACDNTGT 2328
 Qy 2366 PLAVFSGTGLMTKOLYATAYGEIYMDTNPFIQIIGYHGGLYDPLTKLVHMRDYYDLA 2425
 Db 2329 PLAVFSGTGLMTKOLYATAYGEIYMDTNPFIQIIGYHGGLYDPLTKLVHMRDYYDLA 2388
 Qy 2426 GRWTSFDELHMKHLSNNV--PFLYMPKNNPISNSQDIKCFMTDVNSWLLTFFGOLH 2483
 Db 2389 GRWTSFDELHMKHLSNNV--PFLYMPKNNPISNSQDIKCFMTDVNSWLLTFFGOLH 2445
 Qy 2484 NVIPGYPKEDMDAMEPSYELIHTQMTQEDWNSKSLGVQCEVOKAKAFVTLERFDOLY 2543
 Db 2446 NVIPGYPKEDMDAMEPSYELIHTQMTQEDWNSKSLGVQCEVOKAKAFVTLERFDOLY 2505
 Qy 2544 GSTITSCQAPKTKKFASGSGVFGKVFALKDGRVTTDIISVANEDGRRVAAIILNHAH 2603
 Db 2506 QYNEGRCLEGGKQPFAPVSPVFGKVFALKDGRVTTDIISVANEDGRRVAAIILNHAH 2565
 Qy 2604 LENLHFTIDGVTHYFVPGPSEGLAILGSGGRTLENGVNVTVSQINVLNGRTRY 2663
 Db 2566 LENLHFTIDGVTHYFVPGPSEGLAILGSGGRTLENGVNVTVSQINVLNGRTRY 2625
 Qy 2664 TDIOLOQALCLNTRYGTTLDEKARVLELQARAVQAWARQOORLEGBEGLEAWTEG 2723
 Db 2626 ADIOLOQALCLNTRYGTTLDEKARVLELQARAVQAWARQOORLEGBEGLEAWTEG 2685
 Qy 2724 EQQVLSSTGRVGYDGFVVISVEQYBELSDSANNIHFMRQSEMGRR 2769
 Db 2686 EQQVLSSTGRVGYDGFVVISVEQYBELSDSANNIHFMRQSEMGRR 2731

Db 1480 YITDEKKINLRQVTTNGEICILLAGASDCCKNDVNCICYSGGDDAYATDAILNSPSS 1539
Qy 1542 LAVCADGELVADLGNIRIRIFIRKKNPFLNTQNNYELSSPIDOELVDFDTGKHLXTQSL 1601
Db 1540 LAVAPDGTIIADLGNIRIRIRAVSKNKPVLNAFQYEAASPEQELVFNADGHIQHTVSL 1599
Qy 1602 PTGDYLYNFTYTGDDITLTDNNGNMVNRDSTGMPLWLVPDQGVVWVTGNTSALK 1661
Db 1600 VTGEYLYNFTYADNDVELIDNNGNSLKRDSGMPRHLLPNDQIIITLTVTGNGGLK 1659
Qy 1662 SVTQGHELAMWTHGNSGLLAKNSNENGWTTTVEYDSPGLRNTVNTFPTQGVSSFRSDTD 1721
Db 1660 AVSTQNLEGLMTYDGNLTGLLAKSDDETGMTTFYDHDHGLRNTVNRPTGVVTSLSHREME 1719
Qy 1722 SSVHVQVETSSKD-DVITITNLSAGAFYTLLOQVENSYYIIGADGSLRLLLANGMEVAL 1780
Db 1720 KSIITVDIENSNRDNDVTIINLSVEASYIVVQDVENSYQLCSNGTLRWYANGVSP 1779
Qy 1781 QTEPHLAGTVNPTVGRKNTVLPIDNGLNLVWRQKEQARGQVTVFGRRLRVHNRNLLS 1840
Db 1780 HSEPHVLAGTLTPTIGRCNTSLPMENGLNSIWLRLKEQIKGKVTIFGRKLRVHGRNLLS 1839
Qy 1841 LDPDRVTRTEKIYDHRKFTLRILYDQAGRPISLWSPSSRLNGVNVTVSPGGYIAGIQGI 1900
Db 1840 IDYDRNTRTEKIYDHRKFTLRILYDQVGRPFLWLPSSGLAANVSVFFNGLRGLORGA 1899
Qy 1901 MSERMEYDQGRITSRIPADKWTWSTYLEKSWVLLHLSQRYIFBFDKXNDRSSVTMPN 1960
Db 1900 MSERTDIDKQGRIVSRMFADGKWSYSLDKSWVLLQSQRYIFEYDSSDRHLFAVTMPS 1959
Qy 1961 VAROTLETISRGVYRNIYPPENASVIOFDTEGHLLHTFVLGTGRVIVYKYGKLSKL 2020
Db 1960 VARHSMTHSTIGIRNIYNPPESNASVIFDYDDGRILKITSFLGTGRGVYKYGKLSKL 2019
Qy 2021 AETLYDTTKVSFTYDETAGMKITINLQNEGFTCTIRYRQIGLPIIDRQIFRFTBEQVMNAR 2080
Db 2020 SEIYVDSTAVTFYDEITGVLMKNVLSQGGFSCITIRYRKVGPLVDKQIVYRFSSEGMNAR 2079
Qy 2081 PDNYY-DNSPRVTSQWAVINETPLPIDLYRYDVSQKTEQFGKGVYIYDINOIITAVM 2139
Db 2080 FDTYTHDNSPRIASIKVIBETPLVDLYRYDEISGKVEHFGKGVYIYDINOIITAVM 2139
Qy 2140 THTKHFDAVGRMEQVQVEIFRSRLMYMTVOYDNNMGRVVKELVGPVANTRYSYEVDAD 2199
Db 2140 TLSKGFTHGRKEVQVEMFRSLMYMTVOYDSMGVIRKELKGLPYANTTKTYIYDGD 2199
Qy 2200 GQQTQVSNKPLWRYSYDLNGLNHLSPGNSARLTPRYDIRITRLGDVQYKWDDEG 2259
Db 2200 GQLQSAVANDRPTWRYSYDLNGLNHLNPGNSARLMPRLYDRITRLGDVQYKWDG 2259
Qy 2260 FLRQGGDIPEYNSAGLLIKAYNEAGSWSYRYVDGLGRVSSKSHHLQFFYADLTN 2319
Db 2260 YLCQSGSDIPEYNSKGLLTRYAINKASGWSVOYRYDGVSRASRYKTNLGHHLQFYSDLHH 2319
Qy 2320 PTKVTHLYNHSSEITSLYLDLQGLHFAMELSSGDEFYIACDNGITPLAVFSGTGLMIKQ 2379
Db 2320 PTRITHVYNSNSEITSLYLDLQGLHFAMESSGSEYVVASDNTGTPLAVFSINGLMIKQ 2379
Qy 2380 ILITAYGEIYDNDPNQIIGYHGGYLDPLTKLVHMGDRDYDLAGRWTSPDHEHLWKHL 2439
Db 2380 LQYTAIGEIIYDNDPNQIIGYHGGYLDPLTKLVHFTQBDYDLAGRWTSPDYTWNRV 2439
Qy 2440 SSSNMVFNLYMFKNNPINSODIKCFMTDVNSWLLTFQFLHNVIIPGPKPDMAMEP 2499
Db 2440 GKPEA-PFNLYMFKNNPINSNEDLNKYVTDVKSWMVGFQFLSNIPGFRKMYFVPP 2498
Qy 2500 SYELIHTOMKTQEWDNKSILGVOCEVQKQKAFVTLERFDQYVGSITITSQOAPKTKK- 2558
Db 2499 PYELSESQAS----ENGQITGVQOQTERHNQAFALF-----GQVISKKLHAGIREKA 2548
Qy 2559 ---PASSGVSFGKGVKAFALDGRVTTDIISVANEDGERVAAILNHAHYLENLHFTIDGVD 2615
Db 2549 GHWFATPIIGKIMFAIREGRVTTGVSSIASDSRKVASVLNNAYILDKMHSYIEGKD 2608

Qy 2616 THYFVKPQSPSGDLAILGLSGRRRTLENGVNVTVSQINTVNGRRTRRTDIQLOYGALCL 2675
Db 2609 THYFKVIGIAGADGDLVLTGTIGRKVLESGVNVTVSQPTLLVNGRRTRRTNIEFYQYSTLL 2668
Qy 2676 NTRYG---TTLDEEKARVLELARQAVRQAWARBOQRLRBEGBEGLRAWTEGEKQOVLSTG 2732
Db 2669 SIRYGLTPTLDEEAKRVLDQARQALGTAWAKEQKQKARDREGSRLWTEGEKQOVLSTG 2728
Qy 2733 RVQGVGDFGVFVLSVEQYPELSDSANNIHFHROSEMGRR 2769
Db 2729 RVQGVGDFGVFVLEQYPELADSSNIQFLRQNEGKR 2765

RESULT 11

Q9UKZ4
ID Q9UKZ4 PRELIMINARY; PRT; 2725 AA.
AC Q9UKZ4
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Tenascin-M1.
GN TNM1.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20025749; PubMed=10556288;
RA Brandau O., Schuster V., Weiss M., Hellebrand H., Fink F.M.,
RA Kreczy A., Friedrich W., Strahm B., Niemeyer C., Belohradsky B.H.,
RA Meindl A.;
RT "Epstein-Barr virus-negative boys with non-Hodgkin lymphoma are
RT mutated in the SH2D1A gene, as are patients with X-linked
RT lymphoproliferative disease (XLP).";
RL Hum. Mol. Genet. 8:2407-2413 (1999).
DR EMBL; AF100772; AAF04723.1; -
DR HSSP; P00750; ITPG.
DR Genew; HGNC:1117; ODZ1.
DR GO; GO:0005576; C:extracellular; TAS.
DR GO; GO:0008201; F:heparin binding; TAS.
DR GO; GO:0006955; P:immune response; TAS.
DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.
DR GO; GO:0007393; P:neurogenesis; TAS.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR006530; YD.
DR Pfam; PF00008; EGF; 5.
DR Pfam; PF01436; NHL; 6.
DR Pfam; PF05593; RHS_repeat; 6.
DR SMART; SM00181; EGF; 6.
DR TIGRPFAMs; TIGR01643; YD_repeat_2x; 7.
DR PROSITE; PS00022; EGF_1; 8.
DR PROSITE; PS01186; EGF_2; 7.
DR PROSITE; PS00136; SUBTILASE_ASP; 1.
KW EGF-like domain.
SQ SEQUENCE 2725 AA; 304892 MW; 6F383E64C1B2E1B1 CRC64;
Query Match 63.8%; Score 9501; DB 4; Length 2725;
Best Local Similarity 62.0%; Pred. No. 0;
Matches 1737; Conservative 404; Mismatches 550; Indels 112; Gaps 17;
Qy 1 MDVKKRKYRSLTR-RRAERRYTSSSADSBEGKAPQKSYSSSETLKAYDQARLAYGS- 58
Db 1 MEQDCKPYQPLFKVKHEMDLAYTSSDESDGKRQPSYNSRETLHEYNQELRMVNSQ 60
Qy 59 --RVKDIVPQEAEEFRCRTGANFTLRELGLBEVTPPHGTLYRTDI-GLPQCYSYMGAGSD 114
Db 61 SRKRKVEKSTQEMEFCTSTHLC-----SGYQTDHVSRYHGYQLEMGSD 106


```
Db 2206 GEIOYKMEDEGLRQGRNDIFEYNSNGLLOKAYNKASGWTQVYYVYDGLGRRVASKSSLGQ 2265
Qy 2309 HLQFFYADLNPKVTHLYNHSSEITSLYYDLOGLHFLAMELSGSEFFIACDNIGTPLA 2369
Db 2266 HLQFFVDATANP:RVTHLYNHTSEITSLYYDLOGLHFLAMELSGSEFFIACDNIGTPLA 2325
Qy 2369 VFSGTGLMIKOILYIYAYGETYMDNPNFQIIIGYHGLYDPLTKLVHMGRRDDYDLVLAGRW 2428
Db 2326 VFSRGQVKEILYTPYGDYIYHDTYEDFQVILIGFHGGLYDPLTKLVHMGRRDDYDLVLAGRW 2385
Qy 2429 TSPDHEWLKHLSSNW- -PFNLYMFKNPNPINSQDIKCFMTDYNVSNWLLTFQFQLHNV 2486
Db 2386 TTAHYHIIWKOL- -NLLPKFENLYSFENNYFVGKIQDVAKYTTDIRSWLELFGQLHNV 2442
Qy 2487 PGYKPDMDAMESYELIHTQMTQEWDNKSLIGVQCEVQKOLKAFVTLEREDQLYGST 2546
Db 2443 PGFKPPELENELTYELRLQTKQWDPKTLIGQCELOKQLRNFISLQDQPMTPRYN 2502
Qy 2547 ITSQQAPKTKFASGVSFGKGVKFKALDGRVTTDIIISVANEDGRRAAILNHAHYLEN 2606
Db 2503 DGRLEGKQKPRFAAVSFVGKIKFAIKDGIIVTADIIGVANEDSRRLAAILNHAHYLEN 2562
Qy 2607 LHFTIDVDTHYFVKPSPSGDLAILGLSGRITLNGVNVTVSQINTVLNGHTRRYTDI 2666
Db 2563 LHFTIEGRDTHYFKLSLEBDVLIGNTGGRILENGVNVTVSQMTSLNGHTRRFADI 2622
Qy 2667 QLOYGALCLNTRYGTTLDDEKARVLELQARAVRQAWARQOQLRGEGLRAWTEGEKQ 2726
Db 2623 QLOHGALCFNIRYGTTVBEKKNVLEIARQAVQAQWKEORLQSGEGIRAWTEGEKQ 2682
Qy 2727 QVLTGRVQYDGFVLSVQYBELSDSANNIHFMRQSENGRR 2769
Db 2683 QLLSTGRVQYDGFVLSVQYBELSDSANNIHFMRQSEIGRR 2725

RESULT 12
Q9W7R4 PRELIMINARY; PRT; 2590 AA.
AC Q9W7R4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ten-m3.
GN OD3 OR TENM3 OR TEN-M3.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99425191; PubMed=10495292;
RA Mieda M., Kikuchi Y., Hirate Y., Aoki M., Okamoto H.;
RT "Compartmentalized expression of zebrafish ten-m3 and ten-m4,
RT homologues of the drosophila tenm /odd Oz gene, in the central nervous
RT system";
RL Mech. Dev. 87:223-227(1999).
DR EMBL; AB026979; BAA81892.1; -.
DR HSSP; P10968; 7WGA.
DR ZFIN; ZDB-GENE-990714-19; odd3.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR006530; YD.
DR Pfam; PF00008; EGF; 3.
DR Pfam; PF01436; NHL; 5.
DR Pfam; PF05593; Rhs_repeat; 6.
DR PRINTS; PR00011; EGF_LAMININ.
DR SMART; SM00181; EGF; 6.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 7.
DR PROSITE; PS00022; EGF_1; 8.
```

```
DR PROSITE; PS01186; EGF_2; 7.
KW EGF-like domain.
SQ SEQUENCE 2590 AA; 288586 MW; 597592866219148D CRC64;

Query Match 62.5%; Score 9303; DB 13; Length 2590;
Best Local Similarity 63.4%; Pred. No. 0;
Matches 1685; Conservative 384; Mismatches 466; Indels 124; Gaps 19;

Qy 140 RSSCLSSRANSULTITDTEHENTE-----TDHPGQLQHARLRTPPPPLS 184
Db 27 OSSILSNATTQAVQSDSEBEYTAVLYRPVTPQAPSHSCNEQPSNQHQQGGSTLPPVPPP 86
Qy 185 HAHTENQHAASINSLNKNFTPRSNPSPAPTSHLSGEPAGGAQ-----EPAAHQENW 239
Db 87 HKQP-----SVTALNHNLSLRNVSAP-----PAALPAELQTTPESVPLQDSW 132
Qy 240 LLSNINPLETRNLGKQPFGLTLQDNLIEDILGASRHDGAYSDGHFLKFG-GTSPLECT 298
Db 133 VLGSNVPLESR-----HFLKFTGTGTPLFST 159
Qy 299 TSPGYPLTSSTVYSPPPRPLPRSTFARPAFLNKKPSKYNMKCAALSIAIVISATILILA 358
Db 160 ATPGYTMATGAVSPPTPLPNTLSRSAPFKFKSSKYSWRCTALSAMAYILLSVLLC 219
Qy 359 YFVAMHLFGLNHLQPMGQMYE---ITEDTASSMPVPTDVSLYPSGGTGLETPDRKKG 415
Db 220 YCIAMHLFGLNWLQTEGYAFENGQVKSDDTATNAV-----TALSTENK---- 264
Qy 416 TTEGKPSFPFEDSIDSGEIDVGRASQKIPPGTFWESQVFIHPVHLKENVSLGKAAL 475
Db 265 -----VFQENNTIDTGEVDVGRRAVDVPPGTFWRIQLFIQPOSLKFNISVQRAL 317
Qy 476 VGIYGRKGLPPSHTQDFVELLDGRLLTQBARSL---EGTPRQSRGTVPVSSSHETGFIQ 532
Db 318 VGVYGRKGLPPTHQYDFVELLDGRLLTAKEKRGIVEGAARKARSV---NVHEAEFIR 374
Qy 533 YLDSGIWHLAFYNDKSEVVSFLTTAIESVDNCPSNYNGDCISGTCHEGLPLGDDC 592
Db 375 FLDSGTWHLAFYNDGKNAEQVSYNTIITLITECHNCHNGDCKTGTCHEGPGPLGDDC 434
Qy 593 GRACPVLCSGNGQYMKRGCLCHSGWGAECDDVPTNQCIDVACSNHGTCTGTCTCNFGY 652
Db 435 SRACPVLCSGNGQYSRGRCLCYSONKGTCDVPSNQCIDHCSGHGICIMGTCACTGY 494
Qy 653 KGESCEVDVMDPTCSGRGVVRGCHCFVWGGTNCBTPRATCLDQCSGHGTFPDTGL 712
Db 495 KGDNCEVDCLDPSCSHGVCTHGECHCNPGWGNNCBILKTMCPDQCSGHGTYQTESGT 554
Qy 713 CSCDPSWTGDCSIEICAADCGHGVCGVGTCTRCEDGMMGAACDORACHPRCAEHGTCRD 772
Db 555 CTCDINWTGDCSIEVCAVDCSHGVC:IGSGCRCEEGWTGVSVDLKAHPRCTEHTGCKD 614
Qy 773 GKCEGSPGMNGEHCHTIAHYLDRVWKEGCPGLCNGNRCRTLDLNGWHVCVQLGWRGAGCDT 832
Db 615 GKCECHQGTGHECTV-----EGCPGLCNSNRCRTLDQNGWHVCVCPQWRGAGCDV 665
Qy 833 SMETACGSKNDGDGLVDMDPDCCLOPLCHINPLCLGSPNPLDIIEOTQVPSVQQNLH 892
Db 666 AMETLCADKQSEGGDLVDMDPDCCLOSSQOTQFCRGSPPDPFDIISQNPASQQAAQ 725
Qy 893 SFYDRIKFLVGRDSTHIIPGNPFDPGHACVIRGQVMTSDGTPLVGVNISFVNNPLFGYT 952
Db 726 SFYQIISFLTGPSTHVENGENFNRSLVSIIRGQVLTADGTPLIGVNVSVHYPDHGYT 785
Qy 953 ISRQGSFDLVNNGGISIILPERAPPTQHEHTLWLPWDRPFVMTIIMRHEENIPSCD 1012
Db 786 ITQDGMDFILANGGASLTSEFAPPLTFRTVWIPWNVFWMDTLVWKEENIPSCD 845
Qy 1013 LSNRARNPVPVSPSLTSPASSCAEKGPVPEIQALQEEISISGCKMRLSYLSSSTPGVK 1072
Db 846 LSGFIRPSPFLIVATPELSTIFFRSPSPENGIIIPETQVQLQETAIIPGSDIINLMLSSAAGYR 905
Qy 1073 SVLRISLTHTPTTIPFLNMLKVLHMAVEGSLFRKWFAPAAPDLISYFIWDTVDVYNQKVFGLS 1132
```

Db 906 PVLKVTWQTATIPENLKVHLVAVVGRLLPKQKFPAPENLSYTFIWDKTDAYNQRYGLS 965
Qy 1133 EAFVSVGYEYESCPDLLILWEKRTVLOGYBIDASKLWGCWSLDXGHAALNQSGLLHKNGE 1192
Db 966 EAVSVGFYESCLDLILWEKRTAILQGYELDASNMGGWTLDXHVLVDVQNGFLYKNGE 1025
Qy 1193 NQFVSQPPVIGSTMGNGRRSSICSPSCNGLADGNKLLAPVALTCSGSLYVGDENYIR 1252
Db 1026 NVFVSQPPVISTMGNGRRSSICSPSCNGADGNKLLAPVALACGSGSLFVGDFNYIR 1085
Qy 1253 RIPPSSGNVWILELRNKDFRHSHPAHKYYLATDPMGSAVFLSDSNRRVKIKSTVVVK 1312
Db 1086 RIFPSGNVTSMEL-----SNPAHGYLATDPMTGQLYVSDTNSRIRFRPKALTGK 1138
Qy 1313 DLVKNSEVAGTGDOCLPDDTCGGDGKATEATLNPRGITVDKFLYFVDGTMIRI 1372
Db 1139 ELQNAEVAAGTGEQCLPFPDEARCGDGKATEALLGPKGIAVDKNGFIYFVDGTMIRK 1198
Qy 1373 DONGIISTLLGSNDLTSARELSCDSDVMDSIQVLEWPTDLAINPMDNSLYVLDNNVLOI 1432
Db 1199 DRNGIISTLLGSNDLTSARELTCDSNHHIGQVLEWPTDLAINPMDNSLYVLDNNVLOI 1258
Qy 1433 SENHQRIVAGRPHQCOVGDHFLSKVAIHATLESATALAVSHNGVLYIAETDEKKIN 1492
Db 1259 TENQVRIVAGRPHQCOVPGIE-YTMGKRAIQITTEGATAISUSYSGVLYIAETDEKKIN 1317
Qy 1493 RIRQVTSGBISLVAGAPSCDCKNANCDPFGDGDYAKADKINLPSSILAVCADGELYV 1552
Db 1318 RIRQVSTDEISHLAGAPSCDCKNANCDYCGDYAKAELNAPSSILAVSPDGLV 1377
Qy 1553 ADLGNIRIRIRKPKPLNTONMYELSSPIDQELYLPTTGKHLYTQSLPTGDYLYNFY 1612
Db 1378 ADLGNIRIRAIRNRPPQSGSGFVSPASQELYVFDNGTHQYTMSTVTDGKYNFY 1437
Qy 1613 TGDCDITLITDNNGMVNVARDSTGMLVLPDGVVYVMTGNSALKSVTTQGHELAM 1672
Db 1438 SNEDDVTAVTDSGNILVARDPNRPVRIAPDNQVILWITGTNGGLKTLTAQGGELV 1497
Qy 1673 MYHNSGLLATKSNENGWTFTEYDSFGRLNVTPTQGVSSFRSDTSSVHVQVETSS 1732
Db 1498 FTYHNSGLLATKSIQIGMTTFDYDSEGLTNVTPTGTITSLIGEMDRALTVDIETSG 1557
Qy 1733 K-DDVITITNLASGAFYTLLOQVNSYYIGADGSLILLQANGMEVALQTEPHLLAGTV 1791
Db 1558 RDDVSIITNLSSIDSFTYLLVQQLRNSYQVYDNSNRVIYANGMDSHFQTEPHLLAGAS 1617
Qy 1792 NPTVGRNVTLPIDNGLNLVEMRQKEQARGQVTFVGRBLRVHNRNLLSLDFDRVTRTEK 1851
Db 1618 NPTVARRNMTLPGENGQNLVEMRQKEQNRKVVVFGRLVNGRNLSSVDYDRLSRTEK 1677
Qy 1852 IYDHRKFTLRLYDQAGRLSPSSRLNGVNTYSPGYTAGIQTGIMSERMEYDQAG 1911
Db 1678 IYDHRKFLKIYDASGHFTLWVPSKLSVNLTYSTQVTSLQRTTTERVEYDSOG 1737
Qy 1912 RITSRIFADGKTSYTYLEKSWLLHLSORQYIFEPDKNDRLSSVTMPNVAQTLETIRS 1971
Db 1738 RIVSRFPADAKINSYTYLDSKSWLLHLSORQYIFDYDLHGKQIAITMPSVARGTWTIRS 1797
Qy 1972 VGYRNIYQPEGNASVIOFTEDDGHLLHTFYLGTRRVVIYKGLSKLAETLYDTTKVS 2031
Db 1798 VGYRNIYNPPESNASVTVDYSDGQLLRVAHLGTGRVLYKYRRONKLSLEYLSTSRVS 1857
Qy 2032 FTYDETAGMLKTNLQNEGPTCIRVROIGPLIDROI FRFTEBGMNARFDYNDNSFRV 2091
Db 1858 FTYDETAGVLTNVLGSEGFIQSIRVQIGPLVDROI FRFSEBGMNARFDYNDNSFRV 1917
Qy 2092 TSMQAVINETPLPILYRDVDSGKTEQKFGVYIYDINQIITTAVMTHTKHFDAYGRM 2151
Db 1918 TSMQGVINETPLPILYQFDIDISCKVEQKFGVYIYDINQIISTAVMYTKHFDVHGRI 1977
Qy 2152 KEVQYEI FRSLMYMTVOYDNMGRVVKSLKVGYPANTHRYSEYDADGOLQTVSINDKP 2211

Db 1978 KEIQVEIFRSLMYWITIQYDNMGRVTKREIKIGPFANTTKYGEYDVGOLQTVYLNKRM 2037
Qy 2212 LWRYSYDLNGLNHLHLLSCNSARLTPLBYDTRDRITRGLDGYKXMDDEGFLRQSGDIFEF 2271
Db 2038 MWRYNDLNGNHLNLPNSARLTPLRYDURDRITRGLDGYQYRDEDEGFLRQSGAEIFEY 2097
Qy 2272 NSAGLLIKAYNRAGSWSVRYRYDGLGRVRSKSSHSHHLOFFYADLTNPTKVTHLYNHSS 2331
Db 2098 NSKGLLVRVHKSASGWTIQYRYDGLGRRLASRNSLQHLQFFYADLVNPTRIITHVYNHSS 2157
Qy 2332 SEITSLYYDLOGHLPAMELSSGDEFFIACONIGTFLAVFGSTGLMIKOILYITAYGEYMD 2391
Db 2158 SEITSLYYDLOGHLPAMEISGSEEFYIACONTGTPLAVFSNGLLLKQVQTYAYGEYFD 2217
Qy 2392 TNPQFIIIGVHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHLELWKLHSSNVPFNLYM 2451
Db 2218 SNPQFQVIGPHGGLYDPLTKLHFGGERDYDQAGRWTTPDISTWTRV-GKDPAPFNLYM 2276
Qy 2452 FKNNPISNSODIKCFMTDVSNSWLLTFGFLQHNHVPYGPXPDMDAMPEPSYELIHTQMKTO 2511
Db 2277 FRNNPISKIHEVKEYVYTDVNIWLVTFGFHLHNVIPGFPIPKFDLTQPSLEM----RKSO 2332
Qy 2512 EWDNSKSTLGVQCEVQKOLKAFVTLERFDOLYGSTITSQCOAPKTKKFASSGVFGKGVK 2571
Db 2333 LWDPLSISGVQVQVQVMAKAPLSFERMPHIQUSRRSSREKP-WLWFAVTKSLIGKVM 2391
Qy 2572 FAL-KDGRVTTDIISVANEDGRRVAAIILNHAHYLENLHFTIDGVDTHYFVKPQPSSEGDLA 2630
Db 2392 LAITSKGQVATNALNIANEDCIKVTVLNNAFYLEDLHFTVEGRDTHYFIKTSLPESDLG 2451
Qy 2631 ILGSGGERTLENGVNTVQINTVLNGRTRYTDIOLOYGALCLNTRYGTTLDDEKARY 2690
Db 2452 ALRUTSGRKSLENGVNTVQSITTVNGRTRERADVELQYALALHVRVYGMTDDEKARY 2511
Qy 2691 LELARQAVRAWAROEORLREGEGLRAMEGEKQVLTSTGRVQGVYDGFVVSVEQYPE 2750
Db 2512 LEQARQALSSAWAREQORVRDGEGLVWTEGEKRLSSGKVLGYDGYVVSVEQYPE 2571
Qy 2751 LSDSANNIHFMRQSEMGRR 2769
Db 2572 LADSANNVQLRQSEIGKR 2590

RESULT 13
Q9JULC1 PREIMINARY; PRT; 2346 AA.
AC Q9JULC1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OD23 (Fragment).
GN OD23.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Ben-Zur T., Feige E., Motro B., Wides R.;
RT "The mammalian Odz gene family: Homologs of a Drosophila pair rule
RT gene with expression implying distinct yet overlapping developmental
RT roles."
RL Dev. Biol. 117:107-120(2000).
DR EMBL; AF195418; AAF28316.1; -.
DR HSSP; P01492; 1VNB
DR MGI; MGI:1345183; Odz3.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR MGD; MGD:1345183; Odz3.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001258; NHL.
DR InterPro; IPR006530; YD.
DR Pfam; PF00008; EGF; 3.

QY 2371 SGTGLMIKQILYATYGEIYMDTNPNTQIIIGHYHGLYDPLTKLVNMGRRDYDVLGRWTS 2430
 Db 1954 SSNGLMLKQIYATYGEIYFDSNVDFQLVIGFHGLYDPLTKLIHFGERDYDILAGRWT 2013
 QY 2431 PDHELWKLSSNMPNPNLWFKNNPINSQDIKCFMTDVSLLTFCGLHNVIPGYP 2490
 Db 2014 PDIEIKRI-GKDPAPNLFNPNPASKIHVDKDYIIDVSNLWTFGFLHNAIPGP 2072
 QY 2491 KPMDAMEPSYELIHTQMTQWDSKSLGVQCEVQKQKAFVILERFDQYGSTITS 2550
 Db 2073 VPKFDLTPSYELV---KSQMEDVPPIPGVQVQVQVAKAFSLGKMAEV----- 2120
 QY 2551 QAPKTKK-----FASGSGVFGVGFALDKGVTTDIISVANNEDGRRVAAILNAH 2602
 Db 2121 -QVSRKAGAEQSWLWFAVTKSLIGKVMWLVSGQVQVNLNANEDCICVAAVLNNAF 2179
 QY 2603 YLENLHFTIGVTHYFVKPSPSEGLAIILGSGGERTLENGVNVTVSQTIVLNGRTR 2662
 Db 2180 YLENLHFTISGKTHYFIKTTTPESDLGLTLTSGRKALENGINVTVSQSTIVNGRTR 2239
 QY 2663 YTDILOYGALCNTRYGTTLDDEKARVLELQARAVROAWAREQQLREGEGILPAWTE 2722
 Db 2240 FADVEMQFALALFVRYGTMDEKARILEQARALARAWAREQQRVDEGEARLWTE 2299
 QY 2723 GEKQVLSTGRVQYDGFYVSVYEQYPELSDSANNIHFVRQEMGR 2769
 Db 2300 GEKQVLSAGVQYDGYVLSVEQYPELSDSANNIQLRQSEIGR 2346
 RESULT 14
 Q9ULU2 PRELIMINARY; PRT; 2144 AA.
 ID Q9ULU2
 AC Q9ULU2
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAR-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein KIAA1127 (Fragment).
 GN KIAA1127.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=20039618; PubMed=10574461;
 RA Hirosewa M., Negase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;
 RT "Characterization of cDNA clones selected by the GeneMark analysis
 RL from size-fractionated cDNA libraries from human brain."
 RL DNA Res. 6:329-336(1999).
 DR EMBL; AB032953; BAA86441.2;
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR GO; GO:0005198; F:structural molecule activity; IEA.
 DR InterPro; InterPro1881; EGF Ca.
 DR InterPro; InterPro6209; EGF-like.
 DR InterPro; InterPro6210; IEGF.
 DR InterPro; InterPro2049; Laminin_EGF.
 DR InterPro; InterPro1258; NHL.
 DR InterPro; InterPro6530; YD.
 DR Pfam; Pfam0008; EGF; 4.
 DR Pfam; Pfam1436; NHL; 6.
 DR Pfam; Pfam05593; NHL; repeat; 5.
 DR PRINTS; PRO0011; EGF-LAMININ.
 DR SMART; SMART0181; EGF; 6.
 DR TIGRFAMs; TIGR01643; YD repeat_2x; 5.
 DR PROSITE; PS00022; EGF_1; 6.
 DR PROSITE; PS01185; EGF_2; 5.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 2144 AA; 238607 MW; B:93948001AE46B9 CRC64;

Query Match 54.1%; Score 8059.5; DB 4; Length 2144;
 Best Local Similarity 66.9%; Pred. No. 0;
 Matches 1443; Conservative 334; Mismatches 358; Indels 23; Gaps 9;
 621 AECVPPNQCIDVACSNHGTCTGTCTICNPGYKGESCEEYDCMDPTCSRGVGVCRGCHC 680
 Db 1 AECVPPNQCIDVACSNHGTCTGTCTICNPGYKGESCEEYDCMDPTCSRGVGVCRGCHC 60
 QY 681 FVHGKGTNCETPRATCLDQCSHGHTFLPDTGLCSDPSWTHGDCSIEICAADGGHGVCV 740
 Db 61 SPGWGGLNCELARVQCQDQCSHGHTFLPDTGLCSDPSWTHGDCSIEICAADGGHGVCV 120
 QY 741 GGTCTCRDGGWGAACDORACHPRCAEHTCRDCKGCEKSPGWNGEHCTIAHYLDRVVKEGC 800
 Db 121 GGACCEGEGTGAACDORACHPRCAEHTCRDCKGCEKSPGWNGEHCTIAHYLDRVVKEGC 180
 QY 801 PGLCNGRGTCLDLNGWHCVQGLGWRGAGCDTSMETACGDSKNDGDGLVDCMDPDCCLQ 860
 Db 181 PDLCNGRGTCLGQNSQVCQCTGWRGPGCNVAMETSCADNKNDEGDLVDCMDPDCCLQ 240
 QY 861 PLCHINPLCLGSPNPLDIIQETQVPSQNLHGFYDRIKELVGRDSTHIIPGENPPDGGH 920
 Db 241 SACQNSLLCRGRDPLDIIQGGQ--TDWPAVKSFYDRIKELVGRDSTHIIPGENPNSSL 238
 QY 921 ACVIRGQVMTSDGTPLVGNISFVNNPLFGYTTISRDGSGFDLVNNGISILRPERAPFI 980
 Db 299 VSLIRGQVMTSDGTPLVGNISFVNNPLFGYTTISRDGSGFDLVNNGISILRPERAPFI 358
 QY 981 TOEHTLWLPWDRPFVMEITIMRHEENIEPSCDLSNFARPNPVSPSLTSAFSSAEKGP 1040
 Db 359 SQERTVWLPWNSFYAMDTLVNKTENIEPSCDLSNFARPNPVSPSLTSAFSSAEKGP 418
 QY 1041 IVPEIQALQEBISISGCKMRLSYLSSRTPGYKSVLRISLTHPTIPFNLMKVLHLMVAVEGR 1100
 Db 419 IVPEIQALQEBISISGCKMRLSYLSSRTPGYKSVLRISLTHPTIPFNLMKVLHLMVAVEGR 478
 QY 1101 LFRKFAAAPPDLSYFYFWDKTDVYNQKVGLESAFVSVGVYSCPDLLIWEKRTVVLQ 1160
 Db 479 LFKSFQASPNLAYTFIWDKTDVYNQKVGLESAFVSVGVYSCPDLLIWEKRTVVLQ 538
 QY 1161 YEIDASKLGGWSLDKHALNIQSGLHKNGENQFVSQQPFVIGSIMGNGRRSISCPSC 1220
 Db 539 FELDPSNLGGWSLDKHALNIQSGLHKNGENQFVSQQPFVIGSIMGNGRRSISCPSC 598
 QY 1221 NGLADGNKLLAPVALTCSDGSLYVGFNFYIRIFPSGNVTNILELRNKFHSHSFAHK 1280
 Db 599 NGLADGNKLLAPVALTCSDGSLYVGFNFYIRIFPSGNVTNILELRNKFHSHSFAHK 658
 QY 1281 YLLATDPMGAVFLSDNSRVRFKIKSVVVKDIAVKNSEVAVAGTGDCDLPDTRCGDGG 1340
 Db 659 YLLATDPMGAVFLSDNSRVRFKIKSVVVKDIAVKNSEVAVAGTGDCDLPDTRCGDGG 718
 QY 1341 KATEATLTPRGITVDFKGLTYFVDGTWIRRIDQNGIISTLLGSNDLTSARPLSCDSVMD 1400
 Db 719 KATEATLTPRGITVDFKGLTYFVDGTWIRRIDQNGIISTLLGSNDLTSARPLSCDSVMD 778
 QY 1401 ISQVRLWPTDLAINPMDNSLYLDNNVLIQISENHQVIRVAGHPMHQVPGIDHFLLSK 1460
 Db 779 VAQVRLWPTDLAINPMDNSLYLDNNVLIQISENHQVIRVAGHPMHQVPGIDHFLLSK 837
 QY 1461 VAIHATLESATALAVSHNGVLYIAETDEKKINIRIQVTTSGEISLVAGAPSGDCCKNDAN 1520
 Db 838 LAIHALESASAIASHTGVLYITETDEKKINLRQVTTNGEICLLAGAASDCCKNDAN 897
 QY 1521 CDCPSGDDGYAKADKLTPPSLAVCAGDELYVADLGNIRIRFKNKPFLNTQNYELSS 1580
 Db 898 CNYSGDDAYATDAILNSPSSLAVAPDGTIYIADLGNIRIRAVSKNPFVNAQVFAAS 957
 QY 1581 PIDOEYLFDTTGHKLYTQSLPTGDLVNFYTTGDDITLITDNGNNAVNRDSTGMPL 1640
 Db 958 PGEQELVFNADGHIHQVTSVLTGEYLYNFYTTDNDVTELDNNGNSLAKIRDSSGMPR 1017
 QY 1641 WLVPVPGQVYVWVTGNTSALKSVTTQGHELAMMTYHGNGLATKSNENGTTFYEYDSF 1700


```

Db      1018 HLLMPDQIITLVTNGGKLVVSTQNLGLMTYDGNLTGLATKSDGTGWTTFYDYDHE 1077
QY      1701 GRLTNVTFTQCVSSFRSDTSSVHVQVETSSK-DVTTITNLSASGAFYTLIQDVRNS 1759
Db      1078 GRLTNVTFTQCVSSFRSDTSSVHVQVETSSK-DVTTITNLSASGAFYTLIQDVRNS 1137
QY      1760 YTGADSGILLANGMEVALQTEPHLLAGTVPVTKRNVLPIDINGLNLVWRORKEQ 1819
Db      1138 YQLCNNGTLLRVANGMGISFHSBPHVLAGTITFTGRCNISLPMENGLINSIWRURKEQ 1197
QY      1820 ARGQVTFGRRLRVHNNLLSLDFDVRTRTEKLYDDHKKFTLRILYDQGRPLWSPPSR 1879
Db      1198 IKGKVTIFGRLRVHNNLLSLDFDVRTRTEKLYDDHKKFTLRILYDQGRPLWSPPSR 1257
QY      1880 LNVNVTYSPGGYIAGIQGIMSERMEYDQAGITRIPADGKTWSTYLLKSMVLLHS 1939
Db      1258 LAANVSFFNGRLAGLQRCAMSERITDQKGRIVSRMFADGKVMYSYLDKSMVLLQS 1317
QY      1940 QROVIFEDKNDRLSSVTMPNVARQTLTIRSVGYVYRNIVQPEGNASVIOQTEDEGHL 1999
Db      1318 QROVIFEDKNDRLSSVTMPNVARQTLTIRSVGYVYRNIVQPEGNASVIOQTEDEGHL 1377
QY      2000 HTFVLGTGRVVIYKGLSKLAETLYDTTKVSFTYDTAGMLTKINLQNEGFTCTIRYRQ 2059
Db      1378 KTSFLGTGRVVIYKGLSKLAETLYDTTKVSFTYDTAGMLTKINLQNEGFTCTIRYRQ 1437
QY      2060 IGPLIDQIIFRFTBEGMVARPDYNY-DNSFRVTSQAVINETPLDILYRYDDVSGKTE 2118
Db      1438 IGPLVDQIIFRFTBEGMVARPDYNY-DNSFRVTSQAVINETPLDILYRYDDVSGKTE 1497
QY      2119 QFGKFGVIYDINQIITAVMTTKHPDAYGRMKVQVYEFRLMTVMVTVOYDNMGVRVK 2178
Db      1498 HFGKFGVIYDINQIITAVMTTKHPDAYGRMKVQVYEFRLMTVMVTVOYDNMGVRVK 1557
QY      2179 KELKVGYPANTYRYSVEYDAGLOQTVSINDKPLWRYSDYDNGNHLSPGNSARITPLR 2238
Db      1558 RELKLGYPANTYRYSVEYDAGLOQTVSINDKPLWRYSDYDNGNHLSPGNSARITPLR 1617
QY      2239 YDRDRITRLGDVOYKVDGDFRQRGDIFENYSAGLLIKAYNRAGSSVRYRYDGLGR 2298
Db      1618 YDRDRITRLGDVOYKVDGDFRQRGDIFENYSAGLLIKAYNRAGSSVRYRYDGLGR 1677
QY      2299 RVSSKSHSHLOFFYADLTNPTKVTLYNHSSEITSLYDILQGLHFAVELSSGDEFYI 2358
Db      1678 RASVKTNLGHLQVYFSDLNPNTRITHVYNSHSEITSLYDILQGLHFAVELSSGDEFYI 1737
QY      2359 ACNIGTFLAVSGTGLMIKQILYTAGETIYMDTNENFQIIIGYHGGLYDPLTKLYMGR 2418
Db      1738 ASDNTGTPLAVFSINGPMIKQLQYTAGETIYMDTNENFQIIIGYHGGLYDPLTKLYMGR 1797
QY      2419 RDYDVLAGRWTSPDHELKHLSSNVNMPNLYMFKNNPISNSQDIKCFMTDYNMMLTF 2478
Db      1798 RDYDVLAGRWTSPDHELKHLSSNVNMPNLYMFKNNPISNSQDIKCFMTDYNMMLTF 1856
QY      2479 GFQHLNVIPGPKPDMAMSPVELIHTQMTQOEWNSKSIILGVQCEVQKQLKAFVTLER 2538
Db      1857 GFQHLNVIPGPKPDMAMSPVELIHTQMTQOEWNSKSIILGVQCEVQKQLKAFVTLER 1911
QY      2539 FDQLYGTITSCQOAPTKK----FASGSGVFGKGVKFKLQGRVTTDIIISVANEGRV 2594
Db      1912 FDQLYGTITSCQOAPTKK----FASGSGVFGKGVKFKLQGRVTTDIIISVANEGRV 1966
QY      2595 AAILNHARYLENLHFTIDGVDTHFYKPGPSEGDLAILGLSGGRTLENGVNVTVSQINT 2654
Db      1967 ASVLNNAVYLDKWHYSIEGDKTHFYKPGPSEGDLAILGLSGGRTLENGVNVTVSQINT 2026
QY      2655 VLNGRTYRTDIOYQALCLNTRYG---TTLDEKARVLELQRAVQAWAREQRLR 2711
Db      2027 LVNGRTYRTDIOYQALCLNTRYG---TTLDEKARVLELQRAVQAWAREQRLR 2086
QY      2712 EGEGLRAWTEGEKQVLTSTGRVQYDGFVVISVQVPELSDGANNIHFWRQSEMR 2769

```

```

Db      2087 DREGSRWTEGEKQVLTSTGRVQYEGYVLPVEQYFELADSSNTQFLQENMGKR 2144
RESULT 15
Q80TD2 PRELIMINARY; PRT: 1828 AA.
AC Q80TD2;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MKIAA1455 protein (fragment).
GN MKIAA1455.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22579291; PubMed=12639553;
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,
RA Nakajima D., Nagase T., Ohara O., Koga H.;
RT "Prediction of the coding sequences of mouse homologues of KIAA gene:
RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous
RT cDNAs identified by screening of terminal sequences of cDNA clones
RT randomly sampled from size-fractionated libraries.";
RL DNA Res. 10:35-48(2003).
DR EMBL; AK122513; BAC65795.1; -.
DR InterPro; IPR001258; NHL.
DR Pfam; PF01436; NHL; 6.
DR TIGRFAMs; TIGR01643; YD_repeat_2x; 5.
FT NON_TER 1
SQ SEQUENCE 1828 AA; 206341 MW; 5C2AAB91DC3DEB8C CRC64;

Query Match 47.4%; Score 7063; DB 11; Length 1828;
Best Local Similarity 69.8%; Pred. No. 0;
Matches 1287; Conservative 269; Mismatches 263; Indels 24; Gaps 6;

QY 936 LVGNISFVNPLFGYTTISRODGSFLVYNGGISITLLFERAPFTTQRTHTLWLPDRFFV 995
Db 1 LVGNVSFLHSEYGTITRQDGMFLVANGASLTFERSPPFTTQRTHTVWIPWNVFV 60
QY 996 METIIMRHEENIPSCDLSNFARNPVVSPSLTSPASSCAEKGPVIEIQAQBEISIS 1055
Db 61 MDTLVKXENDIPSCDLSGFRVPSPIIVSSPSTFFRSPSPDSPIIPETQVLTHTTIP 120
QY 1056 GCKMPLSVLSRTPGYKSVLISLTHPTIPNLKVLHLMVAVGELFRKWFAPAAADLSY 1115
Db 121 GTDLKSLVSGRAAGYKSVLKITMTQAVIPNLKVLHLMVAVGRLFKWFPASPENLAY 180
QY 1116 FTWDXTDVYNQVFLSEAFVSVGYEYESCPDLILWEKRTTVLQYEIDASKLGWSLDK 1175
Db 181 FTWDXTDVYNQVFLSEAFVSVGYEYESCPDLILWEKRTTVLQYEIDASKLGWSLDK 240
QY 1176 HLAHNIQSILHKGNGENQFVSQDPVTVGSMNGRRRSISCPSCNGADGNKLLAPVAL 1235
Db 241 HVLVDVQNGILYKNGENQFVSQDPVTVGSMNGRRRSISCPSCNGADGNKLLAPVAL 300
QY 1236 TCGSDGSLVGVDFNVIRIFPSGNVTNILELRKDFRSHSPAHKYLIATDPMSCAVFLS 1295
Db 301 AGCIDGSLVGVDFNVIRIFPSGNVTNILELRKDFRSHSPAHKYLIATDPMSCAVFLS 360
QY 1296 DNSNRVFKIKSVVVKLVKNSEYVAGTQDCLPFDDTRCGDGKATEATLTNPRGITV 1355
Db 361 DNTNRIYRPSKLTCAKLTNKAEEVAGTGEQCLPFDEARCCDGGKAVEATLMSPKGMAI 420
QY 1356 DKFGLIYVDGTMIRIDONGIISTLGSNDLTSARPLSCDSVMDISOVLEWPTDLAIN 1415
Db 421 DKNGLIYVDGTMIRIDONGIISTLGSNDLTSARPLSCDSVMDISOVLEWPTDLAIN 480
QY 1416 PMDNLSLYLDNNVLIQISENHQRVIVAGRPMHCQVPGIDHFLLSKVATHTALESATALV 1475

```

```

Db 481 PMDSIIVLNNVVLQJTEHRQVRIIAAGRPMHCQVPGVE-YPVGKHAQVQTLESATAIAV 539
Qy 1476 SHNGVLYIAETDEKKINRIQVTTSGEISLIVAGAPSGCDCKNDANDCDFSGDDGYAKDAK 1535
Db 540 SYSGVLIYITDEKKINRIQVTTDGEISLIVAGIPSECDCKNDANDCDFSGDDGYAKDAK 599
Qy 1536 LNTPESSLAACADGELYVADLGNTRIRFIRKQKFLPNTQNYELSSPIDQELYLFDTTGKH 1595
Db 600 LNAPESSLAAPDGFYIADLGNIRIRAVSKNKLPLNSMNFYEVASPTDQELYIFDINGTH 659
Qy 1596 LYTOSLPTGVLNYFTYTGDDITLITDNNGNMNVNVRDSTGMPLMLVVPDGVGVVWTMG 1655
Db 660 QYTVSLVTGVLNYFNSNDNDVTAVDTSNGNTLIRIRDPNRPVRVVSFDNQVIMLTIG 719
Qy 1656 TNSALKSVTTQGHLEAMWTHGNSGLLATKNSNGWTTFYEDSPFGRLNVTPTQOVSS 1715
Db 720 TNGCLKSMTAGLELVLTFTYHNSGLLATKSDTGTTFPDYDSEGLTNVTPTGWTN 779
Qy 1716 FRSDTSSVHVQVETSSK-DDVTITNLSAGAFYTLLODOVENSYYIGADGSLRLLLAN 1774
Db 780 LHGDMKAITVDIESSREDEVSITSLSSIDSFYIMVQDLRNSYQIGYDGLRIFYAS 839
Qy 1775 GMEVALQTEPHLAGTVNPTVGKENTPLPIDNGLNLVEMRQKEQARGOVTVFGRRLRVH 1834
Db 840 GLDSHYQTEPHVLGAGTANPTAKENMTLPGENGONLVEWFRKEQAQGVNVFGRKLVRN 899
Qy 1835 NRULLSLDFDRVTRTEKIYDDHKKFLRLILYDQAGPSLWSPSSRLNGVNVITYSPGGYIA 1894
Db 900 GRULLSVDFDRTTKTEKIYDDHKKFLRLIAYDTSGHPTLWLPSSKLMAVNVITYSSTGQIA 959
Qy 1895 GIORGIMSERMEYDQAGRIISRIFADGKTWSYILEKSMVLLHLSORQYIFFEFDKNDRLS 1954
Db 960 SIQEGTTSEKVDYDSQRIYVSVPFADGKTWSYILEKSMVLLHLSORQYIFEYDMWDRLS 1019
Qy 1955 SVTMENVARQLETIRSVGYRNIYOPPEGNASVIOPTEDGHLLHTFYLGTRRVYIKY 2014
Db 1020 AITMPSVARHTMOTIRSIGYRNIYNPPESNASHIITDYNEEGLLOTAFLGTSRRVLFRY 1079
Qy 2015 GKLSKLAETLYDTTKYSFTYDETAGMLKTINLQNEGFTCTIRYRQIGPLIDRQIPRETEE 2074
Db 1080 BRQTRLSEILYDSTRVSFTYDETAGVLTWNLOSDFICTIRYRQIGPLIDRQIPRESED 1139
Qy 2075 GVNAREFDYNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVLYYDINQII 2134
Db 1140 GVNAREFDYNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVLYYDINQII 1199
Qy 2135 TTAVMTHTEHFDAYGRMKVOYEIRSLMYMVMYVQYDNMGRVVKELKVGYPYANTIRYSY 2194
Db 1200 STAVMTYTRHFDAGHKEIYEFIRSLMYMVTIYQYDNMGRVTKREIKIGPPFANTIKYAY 1259
Qy 2195 EYDADGLOQTVSINDKPLMRYSLDLNGLNHLSPGNSARLTPLYRIDRITRLGLDVQYK 2254
Db 1260 EYDVGQLOQTVLNEKIMWEYNDLGNLHLPSSSARLTPLYRDLRITRLGLDVQYR 1319
Qy 2255 MDEDFLRQRGDIFENAGLIIKAYNRAGSVRYRVDGLGRRVSSKSSHHLQFPY 2314
Db 1320 LDEDFLRQRGTEIFEYSKGLLTRYVSGSGWTVIYRYDGLGRRVSSKSLGQHLQFPY 1379
Qy 2315 ADLTPKTVTHLVNHSSEITSLYDLOQHLPAMELSSGDEFYIACDNIQTPLAVFSGTG 2374
Db 1380 ADLTPTRITHVNHSSSEITSLYDLOQHLPAMEISSGDEFYIASDNTGTPLAVFSSNG 1439
Qy 2375 LMTKQILYAYGEIYMDTNPFIITGYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHE 2434
Db 1440 LMLKIQYAYGEIYFDSNVDFOLVGFHGGLYDPLTKLIHFGERDYDILAGERWTTPDIE 1499
Qy 2435 LWKHLSSNVMPNLVYMKNNPISNQDIKCFMTDVNSWMLTFGQFOLHNVIPGYPKPDM 2494
Db 1500 IWKRI-GKDPAPNLVYFRNNPASKIHVDKYITDVNSWLVTFGPHLHNAIPGFPVPKF 1558
Qy 2495 DAMEPSYELIHTQMTQEWDSKSIILGVOCEVQOLKAFVTLERFDQLYGSTITSQQAP 2554
Db 1559 DLTEPSYELV-----KSQWEDVPPITGVQOOVARQAKAFSLGKMAEV-----QVS 1605

```

```

Qy 2555 KTKK-----PASSGSVFGKGVKFAKDGRRVTTDIISVANEDGRRAAILNHAHYLEN 2606
Db 1606 RKGAGAEQSWLWFATVKSLIGKGYMLAVSQGRVQTNVLNIANEDCIKVAALVNNAFYLEN 1665
Qy 2607 LHFTIDGVDTHTFYVKPGPSEGDILAILGLSGRRPTLENGVNVITYSQINTVLNGRTRRYTDI 2666
Db 1666 LHFTIEGKDTHTFYIKTTTPESDLGTCLTSGRKALENGINVTVSQSTTVVNGRTRRPADV 1725
Qy 2667 QLQYCALCLNTRYGTTLDEEKARVLELARQAVQAWARQQRLEGESEGLRAWTEGEKQ 2726
Db 1726 EMQFGALAHVRYGWTLLDEEKARILEQARQPALARAWARQQRVDRDGEAGRLWTEGEKR 1785
Qy 2727 QVLSTGRVQGDGFVSVBQYPELSDSANNIHPMROSENGRR 2769
Db 1786 QLLSAGKVQGYDGYVLSVEQYPELADSANNIQFLRQSEIGKR 1828

```

Search completed: June 24, 2004, 16:21:35
Job time : 100 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 16:17:33 ; Search time 33 Seconds
(without alignments)
4331.888 Million cell updates/sec

Title: US-10-029-020-14

Perfect score: 14887

Sequence: 1 MDVKERKPYRSLTRRDAER.....ELSDSANNTHFMQSEMGRR 2769

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PTUS.COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2989	20.1	768	3	US-08-891-845-2
2	2989	20.1	768	4	US-09-514-573-2
3	2982	20.0	560	3	US-08-891-845-4
4	2982	20.0	560	4	US-09-514-573-4
5	1575	10.6	501	3	US-08-891-845-10
6	1575	10.6	501	4	US-09-514-573-10
7	649.5	4.4	2200	4	US-09-796-575-2
8	648.5	4.4	2199	5	US-08-793-273C-2
9	648.5	4.4	2199	5	PCT-US95-11684-2
10	641	4.3	1810	4	US-08-793-273C-4
11	641	4.3	1810	5	PCT-US95-11684-4
12	440	3.0	2471	1	US-08-185-432-16
13	440	3.0	2471	1	US-08-083-590A-19
14	440	3.0	2471	3	US-08-532-384-19
15	440	3.0	2471	4	US-08-899-232-1
16	438.5	2.9	2703	1	US-08-185-432-19
17	438.5	2.9	2703	4	US-08-899-232-4
18	429	2.9	2523	1	US-08-185-432-18
19	428	2.9	2523	4	US-08-899-232-3
20	428	2.9	2556	1	US-08-083-590A-20
21	428	2.9	2556	3	US-08-532-384-20
22	423	2.8	1193	2	US-08-400-159-10
23	423	2.8	1193	3	US-08-611-729A-10
24	419.5	2.8	1219	3	US-08-882-046-5
25	411.5	2.8	1036	4	US-08-068-740A-6
26	411.5	2.8	1187	4	US-09-068-740A-7
27	411.5	2.8	1208	4	US-09-199-865-1

28	411.5	2.8	1218	3	US-08-892-046-2	Sequence 2, Appli
29	411.5	2.8	1218	4	US-09-068-740A-11	Sequence 11, Appl
30	409	2.7	2556	1	US-08-185-432-17	Sequence 17, Appl
31	409	2.7	2556	4	US-08-899-232-2	Sequence 2, Appli
32	408.5	2.7	1218	2	US-08-400-159-6	Sequence 6, Appli
33	408.5	2.7	1218	3	US-08-611-729A-6	Sequence 6, Appli
34	408	2.7	1010	3	US-08-882-046-7	Sequence 7, Appli
35	407.5	2.7	1055	3	US-09-214-278-2	Sequence 2, Appli
36	407.5	2.7	1055	4	US-09-855-722-2	Sequence 2, Appli
37	404.5	2.7	1212	3	US-09-214-278-3	Sequence 3, Appli
38	404.5	2.7	1212	4	US-09-855-722-3	Sequence 3, Appli
39	404.5	2.7	1257	3	US-08-611-729A-8	Sequence 8, Appli
40	403.5	2.7	1065	2	US-08-400-159-8	Sequence 5, Appli
41	402.5	2.7	1238	3	US-09-214-278-5	Sequence 8, Appli
42	402.5	2.7	1238	4	US-09-855-722-5	Sequence 5, Appli
43	401.5	2.7	1218	3	US-09-214-278-7	Sequence 7, Appli
44	401.5	2.7	1218	4	US-09-855-722-7	Sequence 7, Appli
45	396.5	2.7	1358	1	US-08-404-665-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-08-891-845-2
; Sequence 2, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Silkowski, Mark
; TITLE OF INVENTION: Gamma-Herregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 768 amino acids
; TYPE: Amino Acid
; TOPOLOGY: linear
US-08-891-845-2

Query Match 20.1%; Score 2989; DB 3; Length 768;
Best Local Similarity 76.5%; Pred. No. 1.2e-208;
Matches 598; Conservative 22; Mismatches 84; Indels 78; Gaps 10;
QY 1 MDVKERKPYRSLTRRDAERRYTSSADSEEGKAFQKYSVSSETLKAYDQDARLAYGSRV 60
Db 1 MDVKERKPYRSLTRRDAERRYTSSADSEEGKAFQKYSVSSETLKAYDQDARLAYGSRV 60

```

QY 61 KDIVPQAEPCRTGANTLRELGLBEVTPPHGTLVYRTDGLPCQGYSMGAGSDADMEAD 120
DB 61 KDIVPQAEPCRTGANTLRELGLBEVTPPHGTLVYRTDGLPCQGYSMGAGSDADMEAD 120
QY 121 TVLSPEHPVRLWGSTSRSSCLSSRANSNLTLTDEHENTETDHPGGLQNHARLRTTP 180
DB 121 TVLSPEHPVRLWGSTSRSSCLSSRANSNLTLTDEHENTETDHPGGLQNHARLRTTP 180
QY 181 PPLSHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGQEPAAHQENWL 240
DB 181 PPLSHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGQEPAAHQENWL 240
QY 241 LNSNIPLETRNLGKQPFGLTQDNLIEMDILGASRHGDAYSDGHFLFKPGGTSPFLCCTS 300
DB 241 LNSNIPLETRNLGKQPFGLTQDNLIEMDILGASRHGDAYSDGHFLFKPGGTSPFLCCTS 300
QY 301 PGYELTSTVYSPPPRPLPRSTFARPAFNLIKPSKYCNWKAALSAIVISATLVILLAYF 360
DB 301 PGYELTSTVYSPPPRPLPRSTFARPAFNLIKPSKYCNWKAALSAIVISATLVILLAYF 360
QY 361 VAMHLFGLNHLQPMQEGOMYEITEDTASSWPVPTDVSILYPSGGTGLTDPDRKGKGTTEG 420
DB 361 VAMHLFGLNHLQPMQEGOMYEITEDTASSWPVPTDVSILYPSGGTGLTDPDRKGKGTTEG 420
QY 421 PSSFPFEDSFIDSGEIDVGRASQKIPPGTFWRSQVFIHPVHLKFNVSIGKAALVGIY 480
DB 421 PSSFPFEDSFIDSGEIDVGRASQKIPPGTFWRSQVFIHPVHLKFNVSIGKAALVGIY 480
QY 481 RKGLPPSHHTQDFVELLDGRLLTQEARSLGEPTRQSRGTVPSSSHETGFIQVLDGSIWH 540
DB 481 RKGLPPSHHTQDFVELLDGRLLTQEARSLGEPTRQSRGTVPSSSHETGFIQVLDGSIWH 540
QY 541 LAFYNDGKESEVSWFLTTAI---ESVDCNCPNCGVNGD-----CISGTCCHCFLGFL---- 588
DB 541 LAFYNDGKESEVSWFLTTAI---ESVDCNCPNCGVNGD-----CISGTCCHCFLGFL---- 588
QY 589 GPDGGRASCP-----VLCGNGQYMKGRCLCHSGWKAECDDVPTNQC 630
DB 589 GPDGGRASCP-----VLCGNGQYMKGRCLCHSGWKAECDDVPTNQC 630
QY 601 GNELNRRKNKPNQIKQKPKSELINKASLADSGEYM-----CKVISKLG 646
DB 601 GNELNRRKNKPNQIKQKPKSELINKASLADSGEYM-----CKVISKLG 646
QY 631 IDVACSNHGTCTITCICNPGYKESCEVEDMDPTCSGRGVGVRCGECHFCVWGVTNCE 690
DB 631 IDVACSNHGTCTITCICNPGYKESCEVEDMDPTCSGRGVGVRCGECHFCVWGVTNCE 690
QY 647 NDSASAN---ITV-----ESNEIITGPASTEGAYVSSSPIRISVSTEGANTS 693
DB 647 NDSASAN---ITV-----ESNEIITGPASTEGAYVSSSPIRISVSTEGANTS 693
QY 691 TPRAT-----CLDQ-----CSGHGTFL-----PDTGLSCDDPSWTHDCSIEIC 729
DB 691 TPRAT-----CLDQ-----CSGHGTFL-----PDTGLSCDDPSWTHDCSIEIC 729
QY 694 SSTSTTGTSHLVKCAEKEKTCVNGGECFMVKDLISNPSRYLCKCPNFTGDRCONYVM 753
DB 694 SSTSTTGTSHLVKCAEKEKTCVNGGECFMVKDLISNPSRYLCKCPNFTGDRCONYVM 753
QY 730 AA 731
DB 730 AA 731
QY 754 AS 755
DB 754 AS 755

```

RESULT 2

```

US-09-514-573-2
; Sequence 2, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Herogulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/514,573
; FILING DATE: 28 FEB 2000
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/891845
; FILING DATE: 10 JULY 1997
; APPLICATION NUMBER: 60/021640
; FILING DATE: 07/12/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee, Wendy M.
; REGISTRATION NUMBER: 40,378
; REFERENCE/DOCKET NUMBER: P1043
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-1994
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 768 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-09-514-573-2

```

```

Query Match 20.1%; Score 2989; DB 4; Length 768;
Best Local Similarity 76.5%; Pred. No. 1.2e-208;
Matches 598; Conservative 22; Mismatches 84; Indels 78; Gaps 10;

QY 1 MDVKEKPYRSLTRRRDAERYTSSSADSEEGKAPQKYSSESSETLKAYDQDARLAYGSRV 60
DB 1 MDVKEKPYRSLTRRRDAERYTSSSADSEEGKAPQKYSSESSETLKAYDQDARLAYGSRV 60
QY 61 KDIVPQAEPCRTGANTLRELGLBEVTPPHGTLVYRTDGLPCQGYSMGAGSDADMEAD 120
DB 61 KDIVPQAEPCRTGANTLRELGLBEVTPPHGTLVYRTDGLPCQGYSMGAGSDADMEAD 120
QY 121 TVLSPEHPVRLWGSTSRSSCLSSRANSNLTLTDEHENTETDHPGGQNHARLRTTP 180
DB 121 TVLSPEHPVRLWGSTSRSSCLSSRANSNLTLTDEHENTETDHPGGQNHARLRTTP 180
QY 181 PPLSHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGQEPAAHQENWL 240
DB 181 PPLSHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDHSLSGEPAGGQEPAAHQENWL 240
QY 241 LNSNIPLETRNLGKQPFGLTQDNLIEMDILGASRHGDAYSDGHFLFKPGGTSPFLCCTS 300
DB 241 LNSNIPLETRNLGKQPFGLTQDNLIEMDILGASRHGDAYSDGHFLFKPGGTSPFLCCTS 300
QY 301 PGYELTSTVYSPPPRPLPRSTFARPAFNLIKPSKYCNWKAALSAIVISATLVILLAYF 360
DB 301 PGYELTSTVYSPPPRPLPRSTFARPAFNLIKPSKYCNWKAALSAIVISATLVILLAYF 360
QY 361 VAMHLFGLNHLQPMQEGOMYEITEDTASSWPVPTDVSILYPSGGTGLTDPDRKGKGTTEG 420
DB 361 VAMHLFGLNHLQPMQEGOMYEITEDTASSWPVPTDVSILYPSGGTGLTDPDRKGKGTTEG 420
QY 421 PSSFPFEDSFIDSGEIDVGRASQKIPPGTFWRSQVFIHPVHLKFNVSIGKAALVGIY 480
DB 421 PSSFPFEDSFIDSGEIDVGRASQKIPPGTFWRSQVFIHPVHLKFNVSIGKAALVGIY 480
QY 481 RKGLPPSHHTQDFVELLDGRLLTQEARSLGEPTRQSRGTVPSSSHETGFIQVLDGSIWH 540
DB 481 RKGLPPSHHTQDFVELLDGRLLTQEARSLGEPTRQSRGTVPSSSHETGFIQVLDGSIWH 540
QY 541 LAFYNDGKESEVSWFLTTAI---ESVDCNCPNCGVNGD-----CISGTCCHCFLGFL---- 588
DB 541 LAFYNDGKESEVSWFLTTAI---ESVDCNCPNCGVNGD-----CISGTCCHCFLGFL---- 588
QY 589 GPDGGRASCP-----VLCGNGQYMKGRCLCHSGWKAECDDVPTNQC 630
DB 589 GPDGGRASCP-----VLCGNGQYMKGRCLCHSGWKAECDDVPTNQC 630
QY 601 GNELNRRKNKPNQIKQKPKSELINKASLADSGEYM-----CKVISKLG 646
DB 601 GNELNRRKNKPNQIKQKPKSELINKASLADSGEYM-----CKVISKLG 646

```

QY 631 IDVACNHGFCITGTCICNPGYKESCEBVDMDPTCSGRGVCVRGECCHCFVGMGTNCE 690
 Db 647 NDSASAN-----ITIV-----ESNEIITGMPASTEGAYVSSSEFIRISVSTEGANTS 693
 QY 691 TPRAT-----CLDQ-----CSGHGTFL-----DUTGLCSDPDSWTGHDQSIIBC 729
 Db 694 SSTSTTGTSHLVKAEKTEFCVNGGECFVWKDLSNFSRYLCKCPNEFTGDRCONYVM 753
 QY 730 AA 731
 Db 754 AS 755

RESULT 3

US-08-891-845-4
 ; Sequence 4, Application US/08891845
 ; Patent No. 6096873
 ; GENERAL INFORMATION:
 ; APPLICANT: Schaefer, Gabriele M.
 ; APPLICANT: Slikowski, Mark
 ; TITLE OF INVENTION: Gamma-Herregulin
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WinPatIn (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/891,845
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/021640
 ; FILING DATE: 07/12/96
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M.
 ; REGISTRATION NUMBER: 40,378
 ; REFERENCE/DOCKET NUMBER: P1043
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 560 amino acids
 ; TYPE: Amino Acid
 ; TOPOLOGY: Linear

US-08-891-845-4

Query Match 20.0%; Score 2982; DB 3; Length 560;
 Best Local Similarity 99.8%; Pred. No. 2.2e-208;
 Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDVKRKPYSLTRRDAERYTSSADSEGAPOKYSSTSLKAYDODARLAYGSRV 60
 Db 1 MDVKRKPYSLTRRDAERYTSSADSEGAPOKYSSTSLKAYDODARLAYGSRV 60
 QY 61 KDIVPQAEFFCFTGANFTLRELGLEVTTPHGTLYRTDGLPQCGYSMGAGSDADMEAD 120
 Db 61 KDIVPQAEFFCFTGANFTLRELGLEVTTPHGTLYRTDGLPQCGYSMGAGSDADMEAD 120
 QY 121 TVLSPEHPVRLWGRSTRSGRSSCLSSFANSNLTLTDEHENTTDHPGQLNHLRTTP 180
 Db 121 TVLSPEHPVRLWGRSTRSGRSSCLSSFANSNLTLTDEHENTTDHPGQLNHLRTTP 180

QY 181 PPLSHAHTPNQCHAAASINSLNRGNFTPRSNPSPAPTDHLSLSEPPAGCAQPAHAQENWL 240
 Db 181 PPLSHAHTPNQCHAAASINSLNRGNFTPRSNPSPAPTDHLSLSEPPAGCAQPAHAQENWL 240
 QY 241 LNSNIPIETRLNGLQKQFLGTLDNLIEMDILGASRHDGAYSDGHFLFKPGGTSPFLCTTS 300
 Db 241 LNSNIPIETRLNGLQKQFLGTLDNLIEMDILGASRHDGAYSDGHFLFKPGGTSPFLCTTS 300
 QY 301 PGYPLTSTSTVSPPPRPLRSTFARPAFNKKPSKYCNWKAALSAIVISATLVILLAYF 360
 Db 301 PGYPLTSTSTVSPPPRPLRSTFARPAFNKKPSKYCNWKAALSAIVISATLVILLAYF 360
 QY 361 VAMHLFGLNWLQPMEGQYIETDTASSWPVPTDVSILYPSGGTGLETPDRKKGTTGK 420
 Db 361 VAMHLFGLNWLQPMEGQYIETDTASSWPVPTDVSILYPSGGTGLETPDRKKGTTGK 420
 QY 421 PSSFFPEDSFIDSGEIDVRRASOKIPPGTWRQVFDIDHPVHLKFNVLGKAALVGIY 480
 Db 421 PSSFFPEDSFIDSGEIDVRRASOKIPPGTWRQVFDIDHPVHLKFNVLGKAALVGIY 480
 QY 481 RKGLPPSHTOFDVELLDGRLLTQEARSLGTPRQSGTVPSSSHETGFTQYLDGSIWH 540
 Db 481 RKGLPPSHTOFDVELLDGRLLTQEARSLGTPRQSGTVPSSSHETGFTQYLDGSIWH 540
 QY 541 LAFYNDKSESEWSFLTTAI 560
 Db 541 LAFYNDKSESEWSFLTTAI 560
 RESULT 4
 US-09-514-573-4
 ; Sequence 4, Application US/09514573
 ; Patent No. 6500941
 ; GENERAL INFORMATION:
 ; APPLICANT: Schaefer, Gabriele M.
 ; APPLICANT: Slikowski, Mark
 ; TITLE OF INVENTION: Gamma-Herregulin
 ; NUMBER OF SEQUENCES: 11
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: WinPatIn (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/514,573
 ; FILING DATE: 28 FEB 2000
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/891845
 ; FILING DATE: 10 JULY 1997
 ; APPLICATION NUMBER: 60/021640
 ; FILING DATE: 07/12/96
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lee, Wendy M.
 ; REGISTRATION NUMBER: 40,378
 ; REFERENCE/DOCKET NUMBER: P1043
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-1994
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 560 amino acids
 ; TYPE: Amino Acid
 ; TOPOLOGY: Linear
 ; US-09-514-573-4

Query Match 20.0%; Score 2982; DB 4; Length 560;
Best Local Similarity 99.8%; Pred. No. 2.2e-208;
Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MDVKERPYSLRRDAERYTSSGADSEKAPOKSYSSSTLKYAYQDARLAYGSRV 60
DB 1 MDVKERPYSLRRDAERYTSSGADSEKAPOKSYSSSTLKYAYQDARLAYGSRV 60
QY 61 KDIVPOAEFFCTGANTFLRELGLREVTPHGTLYRTDGLPQCCYSMGAGSDADMEAD 120
DB 61 KDIVPOAEFFCTGANTFLRELGLREVTPHGTLYRTDGLPQCCYSMGAGSDADMEAD 120
QY 121 TVLSPEHPVLWGRSTRSGSSCLSSRANSNLTLDTEHENTETDHPGGLQNHARLRTPP 180
DB 121 TVLSPEHPVLWGRSTRSGSSCLSSRANSNLTLDTEHENTETDHPGGLQNHARLRTPP 180
QY 181 PPLSHAHTPNQHAASINSNLRGNFTPRSNPSPAPTDHSLSGEPAGGAQEPAAHAQENWL 240
DB 181 PPLSHAHTPNQHAASINSNLRGNFTPRSNPSPAPTDHSLSGEPAGGAQEPAAHAQENWL 240
QY 241 LNSNIPLNGLKQPPGLTQDNLNLEMDILGASRHDGAYSDGHFLFKPGGTSPFLCTTS 300
DB 241 LNSNIPLNGLKQPPGLTQDNLNLEMDILGASRHDGAYSDGHFLFKPGGTSPFLCTTS 300
QY 301 PGYPLTSSTVYSPPPRPLPRSTTARPAFNKPKSKYCNWKCAALSIVISATLVILLAYF 360
DB 301 PGYPLTSSTVYSPPPRPLPRSTTARPAFNKPKSKYCNWKCAALSIVISATLVILLAYF 360
QY 361 VAMHLFGLNHLQPMQCOMYEITEDTASSWPVPTDVSYPGGTGLTDPDRKGKGTTEK 420
DB 361 VAMHLFGLNHLQPMQCOMYEITEDTASSWPVPTDVSYPGGTGLTDPDRKGKGTTEK 420
QY 421 PSSFFPDSFIDSGEIDVGRASQKIPPGTFRSQVFIIDHPVHLKFNVSGLKAAALVGIY 480
DB 421 PSSFFPDSFIDSGEIDVGRASQKIPPGTFRSQVFIIDHPVHLKFNVSGLKAAALVGIY 480
QY 481 RKGLPSSHOTQDFVELLDGRLLTQARSLEGTPROSGRTVPSSSHETGFIQVLDGSIWH 540
DB 481 RKGLPSSHOTQDFVELLDGRLLTQARSLEGTPROSGRTVPSSSHETGFIQVLDGSIWH 540
QY 541 LAFYNDGKESEVVSFLTTAI 560
DB 541 LAFYNDGKESEVVSFLTTAI 560

RESULT 5
US-08-891-845-10
; Sequence 10, Application US/08891845
; Patent No. 6096873
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,845
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/021640

FILING DATE: 07/12/96
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P1043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-08-891-845-10

Query Match 10.6%; Score 1575; DB 3; Length 501;
Best Local Similarity 64.5%; Pred. No. 5.6e-106;
Matches 332; Conservative 22; Mismatches 83; Indels 78; Gaps 10;

QY 268 MDILGASRHDGAYSDGHFLFKPGGTSPFLCTTSPGYPLTSTVYSPPPRPLPRSTFARPA 327
DB 1 MDILGASRHDGAYSDGHFLFKPGGTSPFLCTTSPGYPLTSTVYSPPPRPLPRSTFARPA 60
QY 328 FNLKKPSKYNWKCAALSIVISATLVILLAYFVAMHLFGLNHLQPMQCOMYEITEDTA 387
DB 61 FNLKKPSKYNWKCAALSIVISATLVILLAYFVAMHLFGLNHLQPMQCOMYEITEDTA 120
QY 388 SSWPVPDVSILYPSGGTGLTDPDRKGKGTTEGKPSFFPEDSFIDSGEIDVGRASQKIP 447
DB 121 SSWPVPDVSILYPSGGTGLTDPDRKGKGTTEGKPSFFPEDSFIDSGEIDVGRASQKIP 180
QY 448 PGTFRSQVFIIDHPVHLKFNVSGLKAAALVGIYGRKGLPPSHOTQDFVELLDGRLLTQEA 507
DB 181 PGTFRSQVFIIDHPVHLKFNVSGLKAAALVGIYGRKGLPPSHOTQDFVELLDGRLLTQEA 240
QY 508 RSLGTPRQSGRTVPSSSHETGFIQVLDGSIWHLAFYNDGKESEVVSFLTTAI 564
DB 241 RSLGTPRQSGRTVPSSSHETGFIQVLDGSIWHLAFYNDGKESEVVSFLTTAI 300
QY 565 NCPNVCYNGD-----CISGTCHCFLGL-----GPDGRASCP----- 598
DB 301 EMKQESAAGSKLVLCRTSEYSSSLRFKFNKNLNRKNKPNQIKIKKPKSKSELRLN 360
QY 599 -VLCSGNGQYMKGRCLCHSGWKGAECVPTNQICIDVACSNHGTCTGTCTCINPGYKGCSC 657
DB 361 KASLADSGEYN-----CKVISKLGNDSSAN-----ITIV-----ESN 393
QY 658 EVDMDPTCSGRGVVGRGECHCFVGMGTNCTTPRAI-----CLDQ-----CSG 702
DB 394 EITGMPASTEGAYVSSSPRISVSVEGANTSSSTSTSTTGTSHLVKCAEKETFCVNG 453
QY 703 HGTFI-----PDGLCSDDPSMTGHDSCSIEICAA 731
DB 454 GECFVWDXLNPSPRYLCKPCNEFTGDRCONVYMAS 488

RESULT 6
US-09-514-573-10
; Sequence 10, Application US/09514573
; Patent No. 6500941
; GENERAL INFORMATION:
; APPLICANT: Schaefer, Gabriele M.
; APPLICANT: Sliwowski, Mark
; TITLE OF INVENTION: Gamma-Heregulin
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080

```
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WinPatIn (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/514,573
FILING DATE: 28 FEB 2000
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/891845
FILING DATE: 10 JULY 1997
APPLICATION NUMBER: 60/021640
FILING DATE: 07/12/96
ATTORNEY/AGENT INFORMATION:
NAME: Lee, Wendy M.
REGISTRATION NUMBER: 40,378
REFERENCE/DOCKET NUMBER: P1043
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-1994
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 501 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
US-09-514-573-10

Query Match      10.6%; Score 1575; DB 4; Length 501;
Best Local Similarity 64.5%; Pred. No. 5.6e-106;
Matches 332; Conservative 22; Mismatches 83; Indels 78; Gaps 10;

QY 268 MDILGASHRGAGSDGHFLKPGGTSPLCTTSPGYPLTSTSVYSPPLPSTFARPA 327
Db 1 MDILGASHRGAGSDGHFLKPGGTSPLCTTSPGYPLTSTSVYSPPLPSTFARPA 60

QY 328 FNLLKPKSKYCNWKAALSAIVISATLVILLAYFVAMHLFGLNWHLOPMQMYEITETA 387
Db 61 FNLLKPKSKYCNWKAALSAIVISATLVILLAYFVAMHLFGLNWHLOPMQMYEITETA 120

QY 388 SSWPVPDVSILYPSGGTGLTPDRKGGTTEGKPSFFPDSFIDSGEIDVGRASQKIP 447
Db 121 SSWPVPDVSILYPSGGTGLTPDRKGGTTEGKPSFFPDSFIDSGEIDVGRASQKIP 180

QY 448 PGTFWRQVFDHPVHLKFNVSLGKAALVGIYGRKGLPPSHQTFDFVELLDGRLLTQEA 507
Db 181 PGTFWRQVFDHPVHLKFNVSLGKAALVGIYGRKGLPPSHQTFDFVELLDGRLLTQEA 240

QY 508 RSLEGTQRGRGIVPPSSHETGFTQYLDSDGIWHLAFYNDGKSEVVSFLTTAI---ESVD 564
Db 241 RSLEGTQRGRGIVPPSSHETGFTQYLDSDGIWHLAFYNDGKSEVVSFLTTAI---ESVD 300

QY 565 NCPNVCYNGD-----CISGTCCHFLGLF---GPDGCRASCP----- 598
Db 301 EMKSQESAASKLVLRCETSSVSLRFKFKNGNLRNKNQNIQKPKGSRLIN 360

QY 599 -VLCSGNGQYKGRCLCHSGWKGAECVDPNQCIDVACSNHGCTTGTCTCNPGYKGESC 657
Db 361 KASLADSGEYM-----CKVISKLGNDSASAN---ITIV-----ESN 393

QY 658 BEVDCMDPTCSGRGVCYRGCHCFVCGWGTNCETPRAT-----CLDQ-----CSG 702
Db 394 EITGMPASTAGAVSSSESIRISVSTEGANTSSSTSTGTSHLVKCAEKEKTCVNG 453

QY 703 HGFTL-----PDTGLCSDPFWGTHDSCIEICAA 731
Db 454 GECFMVKDLNPSRYLCKPNEFTGDRCONVWAS 488

RESULT 7
US-09-796-575-2
; Sequence 2, Application US/09796575
```

```
; Patent No. 6632671
; GENERAL INFORMATION:
; APPLICANT: Genesegues, Inc.
; TITLE OF INVENTION: NANOCAPSULE ENCAPSULATION SYSTEM AND METHOD
; FILE REFERENCE: G332.12-0001
; CURRENT APPLICATION NUMBER: US/09/796,575
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/185,282
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 2
; LENGTH: 2200
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-796-575-2
```

```
Query Match      4.4%; Score 649.5; DB 4; Length 2200;
Best Local Similarity 22.3%; Pred. No. 1.7e-37;
Matches 310; Conservative 166; Mismatches 492; Indels 423; Gaps 64;

QY 545 NDGKESEVVSFLTTAIESVDNCPNVCYNGDCISGTCCHFLGFLGPDGCRASCPVLCG-G 603
Db 208 DDGFTGEDCSQA-----CPSCNDQGVKVCVGLCFEGYAGADCSRICPVPCSEE 259

QY 604 NGQYMKGRCLCHSGWKGAECVDPNQCIDVACSNHGCTTGTCTCNPGYKGESCEVDCM 663
Db 260 HGTCVDGLCVCHDGFAGDDCNKPL--CLN-NCYNRGRCVENECVDEGFTGDCSELICP 316

QY 664 DPTCSGRGVCYRGCHCFVCGWGTNCETPRATCLDQCSHGHTFLPDTGLCSDPSWTGHD 723
Db 317 N-DCFDGRGRCINGTCYCEGFTGEDCGKP--TCPHACTQGR--EBOQCVCDEGFAGLD 371

QY 724 CSIEICAADCGHGVCGVGTCTCCEDGWWGAACDORACHPRCAEHGCTCDGKCECSPGWNG 783
Db 372 CSEKCPADCHNRGCVDRGCECCDDGFTGADGELKCPNGCSGHRGVNGQVCVDEGYTG 431

QY 784 EHCT-----TAHYLDVRVKEG-----CPGLCMNGRCRTLDLNGWHCV 820
Db 432 EDCSOLRCPNDCSHRGCV-EGKVCYCEGFGYDCSDMPCNDCHQHGR--VNGM-CV 486

QY 821 COLGWRGAGC-----DTSMETACGSKNDGDGLV--DCWDPPC-----CL--Q 860
Db 487 CDDGYTGEDCDRQCPDCSNRGLCVDCQVCEDGFTGPDCAELSCPNCHGQRCVNGQ 546

QY 861 PLCH---INPLCLGSPNPLDIIOETQVPVSQONLHSFYDRIKFLVGRDSTHIIPGENPFD 917
Db 547 CVYHEGFGKCKEQRCPSCDCHGQRCVDGQCICHEG-----TGLDC----- 589

QY 918 GGHAC-----VIRQVMTSDGTPLVGVNISFVNNPLFGYTTISRGDGFVLTNGGI 968
Db 590 GQHSFCSPCNNGQCVSGRCICNEG--YSGEDCSEVSPK-----DLV----- 630

QY 969 SIILRERAPFITQ--EHTLMLPMD-RFFVMTIIMRHEENEIPSCDLSNFARPNPVSP 1025
Db 631 -----VTEVTEETVNLAWNMNRVTELV-----VYTP 658

QY 1026 SPLTSSFASCAEKGIVPEIQALQBEISISGCKMRLSYLSSRTPGYKSVLRI---THP 1082
Db 659 T-----HEGGLQMQRVPGDQTS-----TIOLEPGVFIYRFAILNKK 700

QY 1083 TIFPNLMKVMHLMVAVGELFRKFAAAPDLSYFFIWDKTDV-----YNOKVFG- 1130
Db 701 SIPIVSARVATYLPAGEGLKFK-----SIKETSVEVENDPLDIAFETWEIIFRNMKDEGE 756

QY 1131 -----LSEAFVSVGYEYESCPDLI-----LWEKRTVLQO---YEI---D 1164
Db 757 ITKSLRRRPETSVRQGTGLAPGOEYEISLHVKNTRGPGKKEVTTTRLDAPSQIEVKDVTD 816

QY 1165 ASKLGW--SLDKHALNIQSGI-----LHKNGENQFVSQPPVIGSIMGNGRR 1213
Db 817 TTALITWFKPLAIEDIGIELTYGIKVDPGDRTTIDLTEDENQY-----SIGNLKP-TEY 869
```


1214 SISPCSCNGLADGNKLLAPVALTCGSDGLSYGVDFNYIRIFPSPGNVTNILELRNKDFRH 1273
 870 EVSLISRRGDMSSNP--AKETFTTGLDAP-----NLRVVSQTDNSIIL-----EWN 915
 1274 SHSPAHKYIATDPMGSAVFLSDNSRRVPKIKSVTVVVDLVRNSEVAGTGDCUPLPDD 1333
 916 GKAAIDSYRIKAPISGG-----DHAEDVDPKSOQA----- 946
 1334 TRCGDGGKATEATLTPRGITVDKFLIYFVDTGTMIRRIDONGIISTLLGSNDLTSARPL 1393
 947 -----TKTTLTGLRPGT--EYGI-----GVSARKEDKSNPATINAAITELTPKDL 991
 1394 SCDSDMDISQVRLEWPTDLA-----IN-----PMDNSLYVLDNNVWLQISE 1434
 992 QVSTAEATS-LTLLWKTPAKFDYRLNLSLPTQWGVQVQLPRNTTSYVLRG---LEPGQ 1047
 1435 NHQVRIVAGRPMHC-----QVPGIDHLLSKVAIHATLESATAL--AVSHNGVL 1481
 1048 EYNVLLTAERGRHKSAPARVASTEQAPELENLTVEVGWDLRLNLTAAQAYEH----- 1103
 1482 YIAETDE--KKINRIRQVTTSGEISLVAGAPSGDCDCKNDANCDGSDGVDYAKDAKLTPS 1540
 1104 FIIQVQEAANKVEARNLTVP--SLRAVDIPGLKAATPYTVTSYIYQY-----RTPV 1155
 1541 SLAVCAGDELYVADLGNIRIRTRKPKPFLNTQWMLSPIDQELYLFTTGKHLTYOS 1600
 1156 LSAAEASTGE--TPNLGEVVVAEYGVWDALKLN-----WT 1186
 1601 LPTGDYLYNTYTGDDGITLITDNNGMNVNRDSTGMPWLVPVDPGVVMTGNTSAL 1660
 1187 APEGAYEYFIQVEADTVEAAQN-----LTPGGLRSTDLPLGLKAAT 1229
 1661 K-SVTTQCHLAMWYTHGNSGLAATKSNENGWTFYEDYDFGRLTNVTPTQGVSSFRSD 1719
 1230 HYITIRGVQDFSTFSLSEVLTEVPDGNLTVEVSDALRLNWTTPDGYDQFT-- 1287
 1720 TDSVHVQVETSSKDDVTITNLSASCAFYLLQDOVRNSYIYAGDGLRLLLANGMEVA 1779
 1288 -----IQVQEA-----DQVEAHNLTVFSGUR-----SMEI- 1313
 1780 LQTEPHLLAGT 1790
 1314 ---PGLRAGT 1320

RESULT 8
 US-08-793-273C-2
 ; Sequence 2, Application us/08793273C
 ; Patent No. 6482410
 ; GENERAL INFORMATION:
 ; APPLICANT: Crossin, Kathryn L.
 ; APPLICANT: Phillips, Greg
 ; APPLICANT: Prieto, Anne L.
 ; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE ATTACHMENT AND
 ; TITLE OF INVENTION: NEURITE OUTGROWTH, AND METHODS OF MAKING SAME
 ; FILE REFERENCE: BSC00225
 ; CURRENT APPLICATION NUMBER: US/08/793,273C
 ; CURRENT FILING DATE: 1997-05-22
 ; PRIOR APPLICATION NUMBER: PCT/US95/11684
 ; PRIOR FILING DATE: 1995-09-14
 ; PRIOR APPLICATION NUMBER: 08/308,359
 ; PRIOR FILING DATE: 1994-09-16
 ; NUMBER OF SEQ ID NOS: 30
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 2199
 ; TYPE: PR
 ; ORGANISM: Homo Sapiens
 ; US-08-793-273C-2

Query Match 4.4%; Score 648.5; DB 4; Length 2199;
 Best Local Similarity 21.4%; Pred. No. 2e-37;
 Matches 280; Conservative 169; Mismatches 474; Indels 385; Gaps 55;

566 CPSNCGNGDCISGTCHCFLQGLPDCGRASCPVLCSGNGQYMKGRCLCHSGWKGAECDC- 624
 314 CPNCFBGRGRINGTCYCEGFTGEDGKPTCPHACHTQRCBEGQCVCEGAGVDCSE 373
 625 --VPTN-----QCID-----VACSNHGTCITGTCICNPGYKGESE 657
 374 KRCPADCHNRGRVCDGRCECDGFTGADCGELKCPNGCSGHRGVNGQCVCEGTYTGEDC 433
 658 BEVDCMDPTCSGRGVGVGRBCHCFVGGGTNCETPRATCLDQSGHGTFFLPDTGLSCSDP 717
 434 SQLRCPN-DCHSRGRVEGKVCQEQGPKGYDCSD--MSCPNDCHQGRVCV--NGMCVDCD 488
 718 SWTGDCSIEICAAACGCGHGVCGGTCTCRCDGMMGAACDQACHPRCAEHGTCRDGKCEC 777
 489 GYTGEDCDRQCPDCSNRGLCVDGQCVCEGFTGPDCAELSCPNDCGGRGVNGQCV 548
 778 SPWNGEHCITAHYLDVVRVEGCPCLNGNRCRTLDLNGHHCVCQLGWRGAGCDTSMETA 837
 549 HEGFMGDKCK-----EQRCPSDCHGQGRV---VDG-QCICHEGFTGLDCG---QHS 592
 838 CGDSKNDGDLVDCMDPDCCLQPLCHINPLCLG-----SPNPLDIIQETQVPVSQNL 891
 593 CPS-----DONLQCVSGRCICNBEGYSGEDCSEVSP-PKDLV---VTEVTEETV 638
 892 HSFYDR-----IKFLVGRDSTHIIIPGENPFDGGHACVIRGOVMTSDGTPLV-----GVN- 940
 639 NLAWDENMRVTEYLWVYTPH-----EGGLEMQFR---VPGDQSTTIIRELEPGVEY 687
 941 ---ISFVNPLFCYTIISRODGSFDLVNTGSGISIIIRFERAPPTIQTHTLWLWDWFFVM 996
 688 FIRVFAILENK-SIPVSARVATYLPAPEG-----LAFKSIK-ETSVEVWDLDIAFET 740
 997 ETIIMRHEENBIPSCDLNFAFPNVVSPSPLSFASCAEKGPIVPEIOALQBEISISG 1056
 741 WEIIFRNKDEGEITKSLRPE-----TSYROTGLAPG-----QEVEIS- 781
 1057 CKMRLSVLSRTPGYKSVLRISLTHPTIPNLKMLVMAVEGRLFRKFAAPADLSYFF 1116
 782 --LHVKNXNTRGKLVKVTTRLDAPS-----QIEVKDVTDTTALITWPKPLAED- 830
 1117 IWDKTVYVQKVGFLSEAFVSGVEYVESCPLILWEKRTTVLQGYEIDASKLGSWLSKH 1176
 831 -----GIELTYGIKDVFG-----DRTTIDLTEDENQYSIGNLKPDE 867
 1177 HALNIQSGILHKGNGENQFVSQQPVVIGSINGNRRRSISCPSCNGLADGNKLLAPALT 1236
 868 YEVSLL--ISRRGD-----MSSNP-----AKETFT 889
 1237 CGSDGLYVGVDFNYIRIFPSPGNVTNILELRNKDFRSHSPAHKYIATDPMGSAVFLSD 1296
 890 TGLDAP-----NLRVVSQTDNSITI-----EWRNGKAAIDSYRIKAPISGG----- 932
 1297 SNRRVFKIKSVVVDLVRNSEVAGTGDCUPLPDDTRCGDGGKATEATLTPRGITVD 1356
 933 -----DHAEDVDPKSOQA-----TKTTLTGLRPGT-- 958
 1357 KFGLLIYFVDTGTMIRRIDONGIISTLLGSNDLTSARPLSCDSVMDISQVRLEWPTDLA--- 1413
 959 EYGI-----GVSARKEDKSNPATINAAITELTPKDLQVSETAETS-LTLLWKTPAKFD 1012
 1414 --IN-----PMDNSLYVLDNNVWLQISENHQVRIVAGRPMHC----- 1448
 1013 RYRLNYSLPTQWGVQVQLPRNTTSYVLRG---LEPGQYVNVLLTAERGRHKSAPARVKA 1069
 1449 --QVPGIDHLLSKVAIHATLESATAL--AVSHNGVLYIAETDE--KKINRIRQVTTSGEI 1503
 1070 TEQAPELENLTVEVGWDLRLNLTAAQAYEH-----FIIQVQEAANKVEARNLTVP-- 1123
 1504 SLVAGAPSGDCDCKNDANCDGSDGVDYAKDAKLNTPTSSSLAVCAGDELYVADLGNIRIRFI 1563
 1124 SLRAVDIPGLKAATPYTVTSYIYQY-----RTPVLSAAEASTGE--TPNLGEVVVAEV 1175

QY 1564 RKKPFLNTQMYELSSPIDOELYLFTTGKHLTQSLPTGDLVYNFTYTGDDITLTD 1623
 Db 1176 GWDALKLN-----WTAPEGAYEYFFIQVQEAADTVAAQ 1208
 QY 1624 NNGMNVRDSTGMPLWLVPDQVYVMTGNTSALK-SVTQGHSLAMMTYHNSGLL 1682
 Db 1209 N-----LTVPGLSRDLPLGLKAATHYITIRGTQFSITPLSVEVL 1251
 QY 1683 ATKSNGWNTTFEYDSFGLRNTVPTTQVSSFRSDTSSVHVQVETSKDDVTITNL 1742
 Db 1252 TEEVPMGNLTVTEVSDALLRLNMTTPDGTDOFT-----IQVEA----- 1292
 QY 1743 SASGAFYLLQDVRNSYITGAGSLRLLLANGMEVALQEPHLLAGT 1790
 Db 1293 -----DQVEBAHNLTVFGLR-----SMEI-----PGLRAGT 1319

RESULT 9
 PCT-US95-11684-2
 ; Sequence 2. Application PC/TUS9511684
 ; GENERAL INFORMATION:
 ; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE
 ; TITLE OF INVENTION: CYTOSTATIN DERIVATIVES THAT STIMULATE
 ; TITLE OF INVENTION: ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING
 ; TITLE OF INVENTION: AND USING SAME
 ; NUMBER OF SEQUENCES: 28
 ; CORRESPONDENCE ADDRESSES:
 ; ADDRESSEE: The Scripps Research Institute, Office of
 ; ADDRESSEE: Patent Counsel
 ; STREET: 10666 North Torrey Pines Road, TPC 8
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US95/11684
 ; FILING DATE: 14-SEP-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 08/308,359
 ; FILING DATE: 16-SEP-1994
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Logan, April C.
 ; REGISTRATION NUMBER: 33,950
 ; REFERENCE/DOCKET NUMBER: BEC0019P
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619-554-2937
 ; TELEFAX: 619-554-6312
 ; INFORMATION FOR SEQ ID NO: 2:
 ; LENGTH: 2199 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 PCT-US95-11684-2

Query Match 4.4%; Score 648.5; DB 5; Length 2199;
 Best Local Similarity 21.4%; Pred. No. 2e-37;
 Matches 280; Conservative 169; Mismatches 474; Indels 385; Gaps 55;

QY 566 CPSNCGNGDCISGTCGFLGFLGPDGCRASCPVLCSGNGYMKGRCLCHSGWKGAECQ- 624
 Db 314 CPNDCFRGRCINGTCVCEBGTGDCGKPTCPHACTQRCGEGCVCDEGAGVDCSE 373
 QY 625 --VPTN-----QCID-----VACSNHGTCITGTCTCNPGYKGESC 657
 Db 374 KRCPADCHNRGRVCVDCGCECDGFTGADCGELKCPNGSCGHRGVNQCVCDEBGTGDC 433

QY 658 EVDGMDPTCSGRGVYRGECHCFVGWGTNCETPRATCLDQCSGHGTFLPDTGLCSDP 717
 Db 434 SOLRCPN-DCHSRGRCEVGEKVCQEGFGYDCSD--MSCPNDCHQHGRCV--NGMCVDD 488
 QY 718 SWTHGDCSIEICAADCGGHGVCVGTGTCRCEDGWMGAACDQACHPRAEHTGTCRDKCEC 777
 Db 489 GYTGEDCDRCQPRDCSNRGLCVGQCVCEDGFTGPDCAELSCPNDCGRCVNGQVCV 548
 QY 778 SPGNHGEHCTTAHYLDVRVKEGCPGLNNGRCCTLDLNGWHCVQLGWRGACDTSMTA 837
 Db 549 HEGFGKDC-----EORCFSDCHGQRC--VDG-QCICHEGFTGLDCG--QHS 592
 QY 838 CGDSKNDGDGLVDCMDPCCLOPLCHINPLCLG-----SPNPLDIIOETQVPSQONL 891
 Db 593 CPS-----DCNNGQCVSGRCICNEGSEDCSEVSP-PADLV--VTEVTEIV 638
 QY 892 HSFYDR-----IKFLVGRDSTHIIIPGENPFDGGHACVIRGQVMTSDGTPLV-----GVN- 940
 Db 639 NLANDNEMRVTEYLVVYTPH-----EGGLEMQFR---VPGDQSTIIIRELEPGVEY 687
 QY 941 ---ISFVNPLFGYTIISQDGSFPLVTNGGISIIIRERAPFITQEHTLPLWDRFFVM 996
 Db 688 FIRVFAILENKK-SIPVSARVATYLPAPG-----LKFPSIK-ETSVEVWDPLDIAET 740
 QY 997 ETIIMRHEENIPSCDLSNFARPNPVPVSPSLTSPASSCAEKGPVPIQALOEISISG 1056
 Db 741 WEIIFRNWKEDEGEITKSLRPE-----TSVROGLAPG-----QYEIS- 781
 QY 1057 CMRLSYLSRTPGYKSVLRISLTHPTIPFNLKMLVAVGRLFRKWFAPAAADLSYF 1116
 Db 782 --LHIVKNTRGPLKRVTTTLTLDAPS-----QIEVKDVTDTTALITWFKPLAID-- 830
 QY 1117 IWDKTDVYNQKVFGLSEAFVSGVEYESCPDLIIWEKRTTVLQGYEIDASKLGGSLDKH 1176
 Db 831 -----GIELTYGKDVPG-----DFTITLDEDEQYISGNLKPDE 867
 QY 1177 HALNTQSGILHKGNGENFVQQPVIQSGIMNGRRRSISCFSCNGLADGNKLAPVALT 1236
 Db 868 YEVSIL--ISRRG-----MSSNP-----AKETFT 889
 QY 1237 CGSGSLYVGDFNYIRIFPSCGNVTNILELANKOPRSHSPAHKYLLATDPMGAVFLSD 1296
 Db 890 TGLDAP-----RNLRRVSQTDNSITL-----EMRNGKAIDSRYIKYVVISGG----- 932
 QY 1297 SNSREVFRIKSTVVVVKDLVKNSVAVAGTGDCQLPFDTRCGDGGKATEATLTPRGITVD 1356
 Db 933 -----DHAEDVDPKSOQA-----TTKTLTGLRPGT-- 958
 QY 1357 KEGLIFVDGTMIRRIDONGIISTLLGSDNLTSAPPLSCDSVMDISQVRLWPDLA--- 1413
 Db 959 EYGI-----GVSARKEDKSNPATINAAATELDTPKDQVSETAETS--LTLWKTFLAKFD 1012
 QY 1414 ---IN-----PMONSLYVLNNVVLQISENHQVRIVAGRPMHC----- 1448
 Db 1013 RYRLNYSLPTGQWGVQLPRNTTSYVLRG---LEPGQEVNVLITAEKGRHKSPPARVKAS 1069
 QY 1449 --QVPGIDHFLSKVAIHATUESATAL--AVSHNGVLVIAETDE--KKINRIRQVTTSGEI 1503
 Db 1070 TEQAPLENLITVEVGWGLRLNMTAAADQAYEH-----FIIQVQEAANKVEAARNLTVP- 1123
 QY 1504 SIIVAGAPSGCDCKNDANDCDFSGDGDYAKDAKANTPSSSLAVCADGLVVDLGNIRIRFI 1563
 Db 1124 SURAVDIPGLKAATPYTVTSIVGIQY-----RTPVLSAEASTGE--TPNLGEVVVAEV 1175
 QY 1564 RKNKPLNTQMYELSSPIDOELYLFTTGKHLTQSLPTGDLVYNFTYTGDDITLTD 1623
 Db 1176 GWDALKLN-----WTAPEGAYEYFFIQVQEAADTVAAQ 1208
 QY 1624 NNGMNVRDSTGMPLWLVPDQVYVMTGNTSALK-SVTQGHSLAMMTYHNSGLL 1682
 Db 1209 N-----LTVPGLSRDLPLGLKAATHYITIRGTQFSITPLSVEVL 1251
 QY 1683 ATKSNGWNTTFEYDSFGLRNTVPTTQVSSFRSDTSSVHVQVETSKDDVTITNL 1742

Db 1252 TEEVPMGNLTVTEVSWDALRLNWTTPDGYDQFT-----IQVEA----- 1292

Qy 1743 SASGAFYLLQDOVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGT 1790

Db 1293 -----DQVEEAHNLTVPGSLR-----SMEI-----PGLRAGT 1319

RESULT 10

US-08-793-273C-4

; Sequence 4, Application US/08793273C

; Patent No. 6482410

; GENERAL INFORMATION:

; APPLICANT: Crossin, Kathryn L.

; APPLICANT: Phillips, Greg

; APPLICANT: Prieto, Anne L.

; TITLE OF INVENTION: CYTOTACTIN DERIVATIVES THAT STIMULATE ATTACHMENT AND

; TITLE OF INVENTION: NEURITE OUTGROWTH, AND METHODS OF MAKING SAME

; FILE REFERENCE: BEC00228

; CURRENT APPLICATION NUMBER: US/08/793.273C

; CURRENT FILING DATE: 1997-05-22

; PRIOR APPLICATION NUMBER: PCT/US95/11684

; PRIOR FILING DATE: 1995-09-14

; PRIOR APPLICATION NUMBER: 08/308,359

; PRIOR FILING DATE: 1994-09-16

; NUMBER OF SEQ ID NOS: 30

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 4

; LENGTH: 1810

; TYPE: PRT

; ORGANISM: Gallus gallus

US-08-793-273C-4

Query Match 4.3%; Score 641; DB 4; Length 1810;

Best Local Similarity 20.4%; Pred. No. 5.1e-37;

Matches 333; Conservative 219; Mismatches 563; Indels 520; Gaps 72;

Qy 543 FYNDKRESEVVSLTIAESVDNCPNSCYNGDCISGTCFLGFLGPGDGRASCPLVLS 602

Db 332 FCEGYTGDCGELT-----CPNNCNGNGCENGCLVCHEGFGVDDCQKRCPTCN 383

Qy 603 GNGYMKGRCLCHSGMKGABCD---VPTNOCIDVACSNHGTCTITGTCICNPGYKGSCEE 659

Db 384 NRGRVDRGVCHGYLGEDCGELRCPND-----CHNRGRCLNGCQVCDGFIGEDCGE 437

Qy 660 VDCMDPTCSRGVCHVGECHFCVWGWTNCETPRATCLDCSCGHGTFPDTGLCSCDPSW 719

Db 438 LRCPN-DCQQRGRCLNGCQCHGFGTDCGELR--CPNDCNSHGRV--NGQCVCEGY 492

Qy 720 TGHDCSTIEICAADCGHGVGVGTCTCRCDGMMGAAACDORACHPRCAEHGTCRDGKCEQSP 779

Db 493 TGEDCGELRCPNDCHNRGRVCHVGECHFCVWGWTNCETPRATCLDCSCGHGTFPDTGLCSCDPSW 552

Qy 780 GNGEHCTIAHYLDVVKEGCPGLNGNGRGTDLNMGHCVCOLGWRGAGCDTSMETACG 839

Db 553 GFTGEDCR-----ERSCPDNCCNVGRV--VEG-RCVCEGYNG----- 587

Qy 840 DSKNDGCGVDCMDPCCLQPLCHNPLCLGSPNPLDIQETQVPSQONLHSPYDRI- 898

Db 588 -----IDGSD-----VSP-----PTGLTVNVTDKTV-----NLEWKENLV 619

Qy 899 -KELV-----GRDSTHILPGE-----NPPDGGHACVIRQGVMTSDGTPLVGVNIF 943

Db 620 NEVLTVVPTSSGGLDLQFTVPGNQTSATHELEPGVEVIRVFAILKNKSI----- 672

Qy 944 VNNPLFYTTISROGSDFLVNGGIIILIPFERAPITQEHLLWLPDRP---FVMEIIL 1000

Db 673 -----PVSARVATVPAPEG-----LKFES-----VRETSVQVEMDPLSIFDGNELV 715

Qy 1001 MRHEENIPSCDL-SNFARNPVPVSPPLTSFASSCAEKGPIVPEIQALQAEISISGCKM 1059

Db 716 FRNMQKDDNGDITSSILKRE-----TSYNQPGIAPG-----QQYNVS---L 754

Qy 1060 RLSYLSRTPGKSVLRISLTHPTIPFNLMKVHLMVAVEGRLPRKWFAPAAAPDLSYYFIWD 1119

Db 755 HIVKNTRGPCLSRVITTKLDAPS-----QIEAK-----D 784

Qy 1120 KTDVYNQVFGLSFAFVSVEYVESCPLDLILWEKRTTVLOGYE-----IDA 1165

Db 785 VTDI-----TALITMSKPLAEIEGIELTYGPKDVPDGRDITDL 822

Qy 1166 SKLGWSLDKHHALNIQSGILHKGNGENQFVSQPPVIGIMGNRRRS-----ISCPSC 1220

Db 823 SE-----DENQY-----SIGNLRPHTEYEVEVTLISR 849

Qy 1221 NGLADGNKLLAPVALTCGSGSLVYVDFNIRIFPPSGNVTNILELRNKKDFRSHSHPAHK 1280

Db 850 RGDMESDPA-----KEVEFTDLAPRLNKRVSQTDNSITL---EMKFESHANIDN 895

Qy 1281 YLLATDPMGSAFVLSDSNSRRVFKIKSTVVVKDLVKNSVAVAGTGDOCLFPDDTRCGDGG 1340

Db 896 YRIFAPISGG-----DHIELTVPKGNOA----- 919

Qy 1341 KATEATLTNPRGITVDRKFLIYFVDGTMIRRIDQNGIISTLLGSDNLTISARPLSCDSWMD 1400

Db 920 -TTRATLTGLRP-----GTEYGIGVTAVRQ-DRESAPATINAGTDLNPKDLEVSOPTE 971

Qy 1401 ISQVRLEWPTDLA-----INPMENSLVLDNNVVLQISENHQVRIY 1441

Db 972 -TTLISLRWRREPVAKFDYRLTYVSPSKKMEIPLVDSTFIRLG---LDAGTEYITISLV 1027

Qy 1442 AGRPMHCQVP-----GIDHFLLSKVVAIHATLESATALAVSHNGV 1480

Db 1028 AEKGRHKSQPTIKGSTEESEPGLNLSVSETGNDGFOLTWTAADGAYE-----NFV 1078

Qy 1481 LYIAETDEKINRIRQVTTSGEISLVAGAPS---GCCKNDANDCCDFSGDDGVAKDAKL 1536

Db 1079 IQVQSDNPE-----ETWNIITVPGQHSVNVTLGKANTPYNVTLYGVIQRYTKPLY 1130

Qy 1537 NTPSLAVACADGELYVADL-----GNIRIRFKNK---PF-----L 1570

Db 1131 VETTGHAPEVGETVSDIITPESFNLSWTNTGDFDAFTIETDLSNLLPEMFENISGNS 1190

Qy 1571 NTQNYELSSPIDQELYFDFTTGKHL-----TQSLPTGDYLYNTYTGDDITL 1620

Db 1191 RTAHISGLSPSTDFIVLYGIS--HGFRTOAISAAATTEAEPEVDNLLVSADATPDGFRLT 1248

Qy 1621 ITDNG-----NMVAVRSDTGMPLVLPV-----DQVYVWVTMTG 1656

Db 1249 WTADGVDFSVLKIRDTKRSD--PLELIVGHERHTDITGLKBTGEYIEIYGVSSGR 1306

Qy 1657 NS-ALKSVTTQGHELAMMYHGN--SGLLATKSNEN---GWTTFYFYDSFGLTNVTFPT 1710

Db 1307 RSQPINSVAT-----TVGSPKGISFSDITENSARVSWTP-----PR 1343

Qy 1711 QVSSFSRSDTSSVHVQVETSSKDDVTITNLSAGAFYLLQDOVRNSYYIGADGSLRL 1770

Db 1344 SRVDSYRVS-----YVPITGGTPNVVTVDGSKTR-----TKLVKLP-----GVDYVNNI 1388

Qy 1771 LLANGMEVALQTEPHLLAGTVNPTVGKRN--VTLPIDNGLNLVWEV-----QRK 1817

Db 1389 ISVKGFE---ESEP--ISGILKTAIDSPSGLVNMTDSEALATWQPAIAAVDNYIVSYS 1443

Qy 1818 EQARQVQT--VFGRRRLVHNRNLLSLDFDRVTRTEKIYDDHRKFTTILYDQA--GRPSL 1873

Db 1444 SEDPEVTVQVSGMNTVE--YDLNGLRPATETYLRLHAVKDAKSETTSTQTGTGLDAFKDL 1502

Qy 1874 WSPSSRLNGVNVTVSPGYIAGIQGIMSERMEYDQAGRITSIFADGKTWSYTYLEKSM 1933

Db 1503 SATEVQSETAVITWRPP-----RAPVTDYLLYVESIDGRVKEVILDPETTSYTLFE--- 1553

Qy 1934 VLLHRSORQYIFEPKNDRLSSVTMPNVARQTLTIRSVGYRYNNIYOPPEGNASVIOQDT 1993

Db 1554 ---LSPSTQYTVKLQALSR---SMRSKMIQTFTVTTTGLLY---PYPKDCSOALLNGEV 1602

Qy 1994 EDGHLHLHTFYLGTCR 2008

Db	1603 TSG--LYTIYLGNDR 1615
 RESULT 11 PCT-US95-11684-4 ; Sequence 4, Application PC/TUS9511684 ; GENERAL INFORMATION: ; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE ; TITLE OF INVENTION: CYTOTOXIC DERIVATIVES THAT STIMULATE ; ATTACHMENT AND NEURITE OUTGROWTH, AND METHODS OF MAKING ; TITLE OF INVENTION: AND USING SAME ; NUMBER OF SEQUENCES: 28 ; CORRESPONDENCE ADDRESS: ; ADDRESSEE: The Scripps Research Institute, Office of ; ADDRESSEE: Patent Counsel ; STREET: 10666 North Torrey Pines Road, TPC 8 ; CITY: La Jolla ; STATE: CA ; COUNTRY: USA ; ZIP: 92037 ; COMPUTER READABLE FORM: ; MEDIUM TYPE: Floppy disk ; COMPUTER: IBM PC compatible ; OPERATING SYSTEM: PC-DOS/MS-DOS ; SOFTWARE: Patent In Release #1.0, Version #1.25 ; CURRENT APPLICATION DATA: ; APPLICATION NUMBER: PCT/US95/11684 ; FILING DATE: 14-SEP-1995 ; CLASSIFICATION: ; PRIOR APPLICATION DATA: ; APPLICATION NUMBER: US 08/308,359 ; FILING DATE: 16-SEP-1994 ; ATTORNEY/AGENT INFORMATION: ; NAME: Logan, April C. ; REGISTRATION NUMBER: 33,950 ; REFERENCE/DOCKET NUMBER: BSC0019P ; TELECOMMUNICATION INFORMATION: ; TELEPHONE: 619-554-2937 ; TELEFAX: 619-554-6312 ; INFORMATION FOR SEQ ID NO: 4: ; SEQUENCE CHARACTERISTICS: ; TYPE: amino acid ; LENGTH: 1810 amino acids ; TOPOLOGY: linear ; MOLECULE TYPE: protein PCT-US95-11684-4	
Query Match 4.3%; Score 641; DB 5; Length 1810; Best Local Similarity 20.4%; Pred. No. 5.1e-37; Matches 333; Conservative 219; Mismatches 563; Indels 520; Gaps 72;	
Qy	543 FYNDKSESWFLTAATESVDNCPNSCYNGDCISGTCHCFLGLPGDGRASCFVLCS 602
Dd	332 FCEBGTGDGCGLT-----CPNNCNGRCENGCLVCHEGFVGDGDSQKRCPTCN 383
Qy	603 GNQYNKGKCLCHSGWKGAEC---VPNQCIDVACSNHGTCTTGTCINPVGKSCBE 659
Dd	384 NRGRCDVGRVCVCEHYLGDGCELCPND-----CHNRGRNCINGQCVCDEFGSDCGE 437
Qy	660 VDCMDPTCSGRGVCVGECHCFVGGTNCETPRATCLDQCSGHGTFLPTGLCSDDPSW 719
Dd	438 LRCPN--DCOQRGRCINGQCECHEGFIGDCCELR--CPNDCNSGRCV--NGQVCDEGY 492
Qy	720 TGHDSCIETCAADCGHGHCVVGGTCRBDGWGMGAACDQACHPCEABHGTCRBGKECSP 779
Dd	493 TGEDGCELRCPNDCNHRGRCVEGRVCVDCNPFMGEDCGELSCPNDCHQGRCVDCRGVCYE 552
Qy	780 GWNGEHCTAHYLDRVVKEGCPGLCNGRCRTLLDLNGWHCVCOLGWRGAGCDISMETACG 839
Dd	553 GFTGEDCR-----ERSCPNDCNNVGR-----VEG-RVCCEBGIMG----- 587
Qy	840 DSKDNDDGLVDMCDPDCCLOQLCHINPLCLGSNPFLDIIOETQVPSQQNLHSFYDRI- 898

Db 1389 ISVKGFE---ESEP--ISGILKTALDSPSGSLVNVNITDSALATWQPAIAAVDNYIVSVS 1443
 Qy 1818 EQARGQVT--VFGRRLRVHRNLLSLDFDRVTRTEKIYDDHRRKFTLRILYDQA--GRPSL 1873
 Db 1444 SEDEPEVTQMSGNTVE-YDLNGLRPATEYTLRVHAKDAQKSETLSTQTTGLDAPKDL 1502
 Qy 1874 WSPSRRLNGVNVTVSPGGYTAGTQIRGIMSEMEYDQAGRITSIFADGKTWSTYILEKSM 1933
 Db 1503 SATEVQSETAVITWRPP-----RAPVTDYLLTYESIDGRVKSVIILDPETTSYLTLE--- 1553
 Qy 1934 VLLIHSORQIFEPDKNDRLSSVTMPNVARQTLLETIRSVGYRNIYQPPGEGNASVIQDET 1993
 Db 1554 ---LSPSTQYTVKLQALSR-----SMRSKMTQTVFTTGLLY-----PYPKDCSALLNGEV 1602
 Qy 1994 EDGHLHTFYIGTR 2008
 Db 1603 TSG--LYTIYINGDR 1615

RESULT 12

US-08-185-432-16
 ; Sequence 16, Application US/08185432
 ; Patent No. 5750652
 ; GENERAL INFORMATION:
 ; APPLICANT: Artavanis-Tsakonas, Spyridon
 ; APPLICANT: Busseau, Isabelle
 ; APPLICANT: Diederich, Robert J.
 ; APPLICANT: Xu, Tian
 ; APPLICANT: Matsuno, Kenji
 ; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND
 ; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS
 ; NUMBER OF SEQUENCES: 23
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/185,432
 ; FILING DATE: 21-JAN-1994
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 7326-006
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2471 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-185-432-16

Query Match 3.0%; Score 440; DB 1; Length 2471;
 Best Local Similarity 30.1%; Pred. No. 3.8e-22;
 Matches 107; Conservative 44; Mismatches 143; Indels 62; Gaps 20;
 Qy 562 SVDNCPSN-CYNGDC---ISG-TCHCFLGFLGPDGCRASCP-----VLCSNGQY 607
 Db 795 NIDECASNPLNQGTCFDDISGYTCHVLPYTGKNCQTVLAPCSPNCPENAAVCKESPNF 854
 Qy 608 MKGRCLCHSGWKAEGCDVPTNQCIDVACSNHGTC--ITGT--CICNPGYKGSCEE--VD 661

Db 855 BSYTCLCAPGQGORCTIDIDECISKPCMHGLCHNTQGSYMCBPPGFSMDCEIDDD 914
 Qy 662 CWDPTCSGRGVCVRG---ECHCFVGMGGTNCETPRATCLDQ-CSGHGTF--LPDTGLCS 714
 Db 915 CLANPCQNGGCMGVNTFSCLLPFGTGDYKQCDMMNECLSEPCNKGGTCSDYVNSYTK 974
 Qy 715 CDPSTWTHDC--SIEICA-ADCGHGVGVGS---TCRCEDGMMAAC---DQACHPR 763
 Db 975 CQAGFDGVHCCENINNECTESSCFNGGTCVCGINSFSLCPVGTGSCFLHEINECSSHP- 1033
 Qy 764 CAEHCCTCRDG---KCCSCSPWNGEHC-TIAHYLDREVVKSGCPLCNGNGRCTLIDLNGWH 818
 Db 1034 CLNCGTGVDTGTCRCPLGYTGKNCQTLNLCGR-----SPCKNKGTCTCVQKKAESQ 1086
 Qy 819 CVCOLGMRGAGCDSMETACGDSKNDGDGLVDCMDPDCCLQPLCHINPLCLGSPN 874
 Db 1087 CLCPSGWAGAYCDVP-NVSCDIAASRG-----VLVEHLCHQSGVCINAGN 1131

RESULT 13

US-08-083-590A-19
 ; Sequence 19, Application US/08083590A
 ; Patent No. 5786158
 ; GENERAL INFORMATION:
 ; APPLICANT: Artavanis-Tsakonas, S. et al.
 ; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
 ; TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
 ; TITLE OF INVENTION: Nucleic Acids
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/083,590A
 ; FILING DATE: 25-JUN-1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 7326-015
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 212 790-9090
 ; TELEFAX: 212 8698864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 19:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2471 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: peptide
 ; US-08-083-590A-19

Query Match 3.0%; Score 440; DB 1; Length 2471;
 Best Local Similarity 30.1%; Pred. No. 3.8e-22;
 Matches 107; Conservative 44; Mismatches 143; Indels 62; Gaps 20;
 Qy 562 SVDNCPSN-CYNGDC---ISG-TCHCFLGFLGPDGCRASCP-----VLCSNGQY 607
 Db 795 NIDECASNPLNQGTCFDDISGYTCHVLPYTGKNCQTVLAPCSPNCPENAAVCKESPNF 854
 Qy 608 MKGRCLCHSGWKAEGCDVPTNQCIDVACSNHGTC--ITGT--CICNPGYKGSCEE--VD 661

Db 855 ESYTCLCAPGQGRCTIDIDECISKPCNMHGLCHNTQGSYMCECPFGSGMDCEDIDD 914
Qy 662 CMDPTCSGRGVVVRG-----EHCFFVWGWTNCETPRATCLDQ-CSGHGTF--LPDTGLCS 714
Db 915 CLANPCQGGSCMDGVNTFSCCLPFGFTGDKCQTDNNECLSEPCCKNGGTCSDYVNSYTK 974
Qy 715 CDPSTWTHDC--SIEICA-ADCGHGVVVG-----TCRCEDGMMGAAC-----DORACHPR 763
Db 975 COAGFDGVHCENNIENECTESSCFNGGTCVDGINSFCLCPVGTGSGFCLHEINECSSHP- 1033
Qy 764 CAEHGTCRDG-----KCECSPGWNGEHC-TIAHYLDRVVKEGCGPLCNGNGRCTLDLNGWH 818
Db 1034 CLNEGTCVDGLGYRCSPLGYTGKNCQTLVNLCSR-----SPCKNKGTCVQKKAESQ 1086
Qy 819 CVCQLGWRGAGCDTSMETACGDSKNDGDLGVDMPDCCLOPLCHINPLCLGSPN 874
Db 1087 CLCPSGWAGAYCDVP-NVSCDIAASRRG-----VLVEHLQHSQGVGINAGN 1131

RESULT 14

US-08-532-384-19
; Sequence 19, Application US/08532384
; Patent No. 6083904
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And
; TITLE OF INVENTION: Nucleic Acids
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/532,384
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/083,590
; FILING DATE: 25-JUN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 7326-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212 790-9090
; TELEFAX: 212 8698864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2471 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-532-384-19

Query Match 3.0%; Score 440; DB 3; Length 2471;
Best Local Similarity 30.1%; Pred. No. 3.8e-22;
Matches 107; Conservative 44; Mismatches 143; Indels 62; Gaps 20;

Qy 562 SVDNCPSPN-CYNGNDC---ISG-TCHCFLGFLGPDCCGRASCP-----VLCSGNGQY 607
Db 795 NIDECASNPLNQGTCTFDDISGYTCHVLPYTGKNCQTLVAPCSNPENAAVCKESPNF 854

Qy 608 MKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTG---ITGT--CICNPGYKGESCEE--VD 661
Db 855 ESYTCLCAPGQGRCTIDIDECISKPCNMHGLCHNTQGSYMCECPFGSGMDCEDIDD 914
Qy 662 CMDPTCSGRGVVVRG-----EHCFFVWGWTNCETPRATCLDQ-CSGHGTF--LPDTGLCS 714
Db 915 CLANPCQGGSCMDGVNTFSCCLPFGFTGDKCQTDNNECLSEPCCKNGGTCSDYVNSYTK 974
Qy 715 CDPSTWTHDC--SIEICA-ADCGHGVVVG-----TCRCEDGMMGAAC-----DORACHPR 763
Db 975 COAGFDGVHCENNIENECTESSCFNGGTCVDGINSFCLCPVGTGSGFCLHEINECSSHP- 1033
Qy 764 CAEHGTCRDG-----KCECSPGWNGEHC-TIAHYLDRVVKEGCGPLCNGNGRCTLDLNGWH 818
Db 1034 CLNEGTCVDGLGYRCSPLGYTGKNCQTLVNLCSR-----SPCKNKGTCVQKKAESQ 1086
Qy 819 CVCQLGWRGAGCDTSMETACGDSKNDGDLGVDMPDCCLOPLCHINPLCLGSPN 874
Db 1087 CLCPSGWAGAYCDVP-NVSCDIAASRRG-----VLVEHLQHSQGVGINAGN 1131

RESULT 15

US-08-899-232-1
; Sequence 1, Application US/08899232
; Patent No. 6436650
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, Spyridon
; APPLICANT: Qi, Huilin
; TITLE OF INVENTION: ACTIVATED FORMS OF NOTCH AND METHODS BASED THEREON
; FILE REFERENCE: 7326-046
; CURRENT APPLICATION NUMBER: US/08/899,232
; CURRENT FILING DATE: 1997-07-23
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2471
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-899-232-1

Query Match 3.0%; Score 440; DB 4; Length 2471;
Best Local Similarity 30.1%; Pred. No. 3.8e-22;
Matches 107; Conservative 44; Mismatches 143; Indels 62; Gaps 20;

Qy 562 SVDNCPSPN-CYNGNDC---ISG-TCHCFLGFLGPDCCGRASCP-----VLCSGNGQY 607
Db 795 NIDECASNPLNQGTCTFDDISGYTCHVLPYTGKNCQTLVAPCSNPENAAVCKESPNF 854
Qy 608 MKGRCLCHSGWKGAECDDVPTNQCIDVACSNHGTG---ITGT--CICNPGYKGESCEE--VD 661
Db 855 ESYTCLCAPGQGRCTIDIDECISKPCNMHGLCHNTQGSYMCECPFGSGMDCEDIDD 914
Qy 662 CMDPTCSGRGVVVRG-----EHCFFVWGWTNCETPRATCLDQ-CSGHGTF--LPDTGLCS 714
Db 915 CLANPCQGGSCMDGVNTFSCCLPFGFTGDKCQTDNNECLSEPCCKNGGTCSDYVNSYTK 974
Qy 715 CDPSTWTHDC--SIEICA-ADCGHGVVVG-----TCRCEDGMMGAAC-----DORACHPR 763
Db 975 COAGFDGVHCENNIENECTESSCFNGGTCVDGINSFCLCPVGTGSGFCLHEINECSSHP- 1033
Qy 764 CAEHGTCRDG-----KCECSPGWNGEHC-TIAHYLDRVVKEGCGPLCNGNGRCTLDLNGWH 818
Db 1034 CLNEGTCVDGLGYRCSPLGYTGKNCQTLVNLCSR-----SPCKNKGTCVQKKAESQ 1086
Qy 819 CVCQLGWRGAGCDTSMETACGDSKNDGDLGVDMPDCCLOPLCHINPLCLGSPN 874
Db 1087 CLCPSGWAGAYCDVP-NVSCDIAASRRG-----VLVEHLQHSQGVGINAGN 1131

Search completed: June 24, 2004, 16:22:59
Job time : 41 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 24, 2004, 16:14:18 ; Search time 24 Seconds
(without alignments)
6007.594 Million cell updates/sec

Title: US-10-029-020-14
Perfect score: 14887
Sequence: 1 MDVXKPYSLRRDAER.....ELSDSANNIHFQSEMGRR 2769

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141691

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	652	4.4	1745	1 TENA_PIG	Q29116 homo sapien
2	649.5	4.4	2201	1 TENA_HUMAN	P24821 homo sapien
3	644	4.3	1808	1 TENA_CHICK	P10039 gallus gall
4	597.5	4.0	4289	1 TENX_HUMAN	P22105 homo sapien
5	440	3.0	2471	1 NTC2_HUMAN	O04721 homo sapien
6	438.5	2.9	2703	1 NTC2_DROME	P07207 drosophila
7	434	2.9	1213	1 JAG3_BRARE	Q90554 brachydanio
8	429	2.9	2524	1 NTC1_XENLA	E21783 xenopus lae
9	428	2.9	2556	1 NTC1_HUMAN	P46531 homo sapien
10	427.5	2.9	2531	1 NTC1_MOUSE	O01705 mus musculus
11	424	2.8	2470	1 NTC2_MOUSE	O35516 mus musculus
12	423.5	2.8	1064	1 FBPI_STRPU	P10079 stronglyloce
13	422.5	2.8	2531	1 NTC1_RAT	O07008 rattus norv
14	420	2.8	2471	1 NTC2_RAT	Q9q30 rattus norv
15	419.5	2.8	1219	1 JAG1_RAT	O63722 rattus norv
16	413	2.8	2437	1 NTC1_BRARE	P46530 brachydanio
17	411.5	2.8	1218	1 JAG1_HUMAN	P78504 homo sapien
18	410.5	2.8	1242	1 JAG1_BRARE	O90557 brachydanio
19	405.5	2.7	1218	1 JAG1_MOUSE	O9qxx0 mus musculus
20	404.5	2.7	1238	1 JAG2_HUMAN	Q9y219 homo sapien
21	400.5	2.7	1295	1 GLPI_CABEL	P13508 caenorhabdi
22	397.5	2.7	2319	1 NTC3_RAT	Q9r172 rattus norv
23	395	2.7	1247	1 JAG2_MOUSE	O9qy65 mus musculus
24	392.5	2.6	2321	1 NTC3_HUMAN	Q9um47 homo sapien
25	391.5	2.6	1202	1 JAG2_RAT	P97607 rattus norv
26	387.5	2.6	1964	1 NTC4_MOUSE	P31695 mus musculus
27	386	2.6	2003	1 NTC4_HUMAN	Q99466 homo sapien
28	384	2.6	2318	1 NTC3_MOUSE	C61982 mus musculus
29	383	2.6	1408	1 SERR_DROME	P18168 drosophila
30	382.5	2.6	833	1 NTC2_MOUSE	P59222 mus musculus
31	382.5	2.6	2139	1 CRB_DROME	P10040 drosophila
32	378.5	2.5	1376	1 CRBH_HUMAN	P82279 homo sapien
33	373.5	2.5	833	1 DL_DROME	P10041 drosophila

34	369.5	2.5	830	1 SREC_HUMAN	Q14162 homo sapien
35	363	2.4	870	1 SRC2_HUMAN	Q96gp6 homo sapien
36	356.5	2.4	473	1 FP2_MVTGA	Q25464 mytilus gal
37	355	2.4	570	1 FBPI_STRPU	P49013 stronglyloce
38	343.5	2.3	1429	1 LI12_CABEL	P14585 caenorhabdi
39	342	2.3	723	1 DLL1_HUMAN	O00548 homo sapien
40	340.5	2.3	686	1 DLL4_MOUSE	Q9j171 mus musculus
41	340	2.3	714	1 DLL1_RAT	P97677 rattus norv
42	337	2.3	1560	1 TENN_MOUSE	Q80271 mus musculus
43	334.5	2.2	379	1 WIF1_HUMAN	Q9y5w5 homo sapien
44	331.5	2.2	379	1 WIF1_MOUSE	Q9wua1 mus musculus
45	327.5	2.2	378	1 WIF1_BRARE	Q9w6f9 brachydanio

ALIGNMENTS

RESULT 1					
TENA_PIG					
ID	TENA_PIG	STANDARD;	PRT;	1746 AA.	
AC	Q29116; P98142;				
DT	15-JUL-1999 (Rel. 38, Created)				
DT	15-JUL-1999 (Rel. 38, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neuronectin)				
DE	(GMEW) (JI) (Microendinous antigen) (Glioma-associated-extracellular				
DE	matrix antigen) (GP 150-225) (Tenascin-C) (TN-C) (P230).				
GN	TNC OR HXB.				
OS	Sus scrofa (Pig).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
OX	NCBI_TaxID=9823;				
RN	[1]_TaxID=9823;				
RP	SEQUENCE FROM N.A. (ISOFORMS MAJOR; MINOR-1 AND MINOR-2).				
RC	TISSUE=Submaxillary gland;				
RX	MEDLINE=92104189; PubMed=1722152;				
RA	Nishi T., Weinstein J., Gillespie W.M., Paulson J.C.;				
RT	"Complete primary structure of porcine tenascin: detection of				
RT	tenascin transcript in adult submaxillary glands.";				
RL	Eur. J. Biochem. 202:643-648(1991).				
RN	[2]				
RP	SEQUENCE OF 813-825; 887-917; 998-1011; 1597-1608 AND 1719-1730.				
RC	TISSUE=Fetal brain;				
RX	MEDLINE=98158323; PubMed=9498558;				
RA	Wakatsuki S., Ho S.H., Arioka M., Yamasaki M., Kitamoto K.;				
RT	"Isolation and characterization of a 230 kDa protein (p230)				
RT	specifically expressed in fetal brains: its involvement in neurite				
RT	outgrowth from rat cerebral cortex neurons grown on monolayer of				
RT	astrocytes.";				
RL	J. Biochem. 122:1146-1152(1997).				
CC	FUNCTION: SAM (substrate-adhesion molecule) that appears to				
CC	inhibit cell migration. May play a role in supporting the growth				
CC	of epithelial tumors. Is a ligand for integrins alpha-8/beta-1,				
CC	alpha-9/beta-1, alpha-v/beta-3 and alpha-v/beta-6.				
CC	FUNCTION: Plays a role during early brain development particularly				
CC	in growth cone guidance. Involved in neurite outgrowth from				
CC	cortical neurons grown on the monolayer of astrocytes.				
CC	SUBUNIT: Hexameric. A homotrimer may be formed in the triple				
CC	coiled-coil region and may be stabilized by disulfide rings at				
CC	both ends. Two of such half-hexabrachions may be disulfide linked				
CC	within the central globule.				
CC	SUBCELLULAR LOCATION: Secreted; extracellular matrix.				
CC	ALTERNATIVE PRODUCTS:				
CC	Event=Alternative splicing; Named isoforms=3;				
CC	Comment=Isoforms are produced in a tissue-and time-specific				
CC	manner during development;				
CC	Name=Minor-2;				
CC	Isoid=Q29116-1; Sequence=Displayed;				
CC	Name=Major;				
CC	Isoid=Q29116-2; Sequence=VSP_001416;				
CC	Name=Minor-1;				
CC	Isoid=Q29116-3; Sequence=VSP_001417;				
CC	TISSUE SPECIFICITY: Submaxillary glands and brain.				

Db 535 DCHGRGRC---VNG-QVCHEGFTGDKQRRCPGDCGRCVQGVQCVQCHGFTGLDCG 590
Qy 854 DPDC-----CLQPLCHINPLCLG-----SPNPLDIIQETQVPVSOQNLSFYDR- 897
Db 591 QRSPCDNCNWGCQVSCRCICNEGSGEDCSQSP-PKDLI---VTEVTEETVNLANDNE 646
Qy 898 ---IKFLVGRDSTH-----IIPGNPFDDGHACVIRQVMTSDGTPLVGNV-----I 941
Db 647 MRVTEYLIVTPTTHDEGLEMQFRVPGD-----QTSTTIRELEPGVEYFIRVF 693
Qy 942 SFVNNPLFGYITSGDGSFELVNGGJSIILRERAPFIQEHITLMPDREFVMTIIM 1001
Db 694 AILENKK-SIPVSARVATY-LPTPEG-----LKFASKIK-ETSVEVEMDPLDIAETWEIIF 746
Qy 1002 RHENEIPSCLDNFARNPNVSPSLTSPASSCAEKGPVIPVIOALQEBISISGCKMRL 1061
Db 747 RNMNKEDEGEITKSLRPE-----TTVRQTGLAPG-----QBEYIS---LHI 785
Qy 1062 SYLSSRTPGYKSVLRISLTHPTTFNLMKVLHVMVAVEGRLFRKWFAPAAAPDLSYFIWDKT 1121
Db 786 VKNTRPGGLKRVTTTDLAPS-----QIEAKDVTDTTALITWFKPLABID----- 831
Qy 1122 DVYNQKVFGLSEAFVSQVGYEYSCPDILMEKETTIVLQGYEIDASKLGGWSLDKHALNI 1181
Db 832 -----GIELTYGKDVPG-----DRTTIDLTHEENQYSGNKLKPTDEYEVSL 873
Qy 1182 QSGILHKGNGENFVSOQPPVIGSIMNGRRRISCPSCNGLADGNKLLAPVALTCGSDG 1241
Db 874 IS-----RR-----ADMGSNPAKETFTTGLDA 895
Qy 1242 SLVGVDFNYIRLFPFSGNVTNILELRNKPFRHSHSPAHKYKYLATDPMGSAVFLSDNSRR 1301
Db 896 P-----RNLRISQTNSTL-----EWRNKAADTVRIKVAISGG----- 933
Qy 1302 VFKIKSVVVKDLVKNSEVAVAGTQCLPFDDTRCGDGGKATATLNPRTGIVTDKFGLI 1361
Db 934 -----DHAEEVPRSPQ-----TTKATLTGLRPGT---EYGI- 963
Qy 1362 YFVDGTWIRIDQNGIISTLLGNDLTSARPLSCDSVMDISQVRLWPTDLA-----IN 1415
Db 964 ---GSAVKGDESDFATNAATDLPDPKDFRVSELKE-SSUTLLWRTPLAKFDRYLN 1018
Qy 1416 -----PMDNSLVLDNNVVLQISENHQVRIVAGRPMHCQVPG----- 1452
Db 1019 YGLPSGQVVEQLPRNATSYLRG---LEPGQVETILLTAEGRHKSKPARVKASTAGEP 1075
Qy 1453 -----IDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRI-ROVTTIS 1500
Db 1076 EIGNLSVSDITPESFSLSWTATEGAFETFTIEIDSNRFL---ETWEYNISGAERTAHIS 1132
Qy 1501 G-----EISLVAGAPSGCCKNDANCDFSGDDGYAKDAKLNTPRSSLAVCAGE----- 1549
Db 1133 GLRPGNDFIVLSGLAFGIQTK-----PISATATTEAEFEVDN 1170
Qy 1550 LYVADLGNIRIFIRKKNKPLNTQNMVELSSPIDOEY-----LFDTTGKHLYTQSLEPT 1603
Db 1171 LLVSDA-----TPDGPRLSWTADGCVDFSVLKIRDTK-----KQSEPL 1209
Qy 1604 GDVLYNFTYGDGDIITLTDNNGMNVNRDSTGMPLWLVPQGVYVWMTGNTNSALKSV 1663
Db 1210 -----EITLLASERTTRDITGLREATEYEI-----ELYGISGKRS----- 1244
Qy 1664 TTQGHELAMTYHGNISGLLTKNGENWTTFYEYDSFGRLTNVFTPTGQVSSFRSDTSS 1723
Db 1245 -----QPVSAIATAMGSPKEITFSDITENSATVSMWVPTAQVESFR-----I 1287
Qy 1724 VHVQVETSSKDDVTITTNLSAGAFYTLLOQOVNSYIIGADGSLRLILLANGNEVALQTE 1783
Db 1288 TYVPITGAPSVVTVDTGKTQTRILLRL-----PGVEYLSVIAVKGFE---ESE 1334
Qy 1784 PHLLAGVWNPVTVGRNVTLPIDNGLNLVENVQRKEQAGQVTVFGRRLRVHNNLLSLDF 1843

RESULT 2

TENA_HUMAN

ID TENA_HUMAN STANDARD; PRT; 2201 AA.
AC P24821; Q14583; Q15567;
DT 01-MAR-1992 (Rel. 21, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neuronectin)
DE (GEMX) (J1) (Microendinous antigen) (Glioma-associated-extracellular
DE matrix antigen) (GP 150-225) (Tenascin-C) (TN-C).
TNC OR HXB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-32.
RC TISSUE=Fetal brain, and Melanoma;
RX MEDLINE=91187670; PubMed=1707164;
RA Siri A., Carnemolla B., Seginati M., Leprini A., Casari G.,
RA Baralle F., Zardi L.;
RA "Human tenascin: primary structure, pre-mRNA splicing patterns and
RT localization of the epitopes recognized by two monoclonal
RT antibodies.";
RL Nucleic Acids Res. 19:525-531(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95155442; PubMed=7531707;
RA Gherzi R., Carnemolla B., Siri A., Ponassi M., Balza E., Zardi L.;
RT "Human tenascin gene. Structure of the 5'-region, identification, and
RT characterization of the transcription regulatory sequences.";
RL J. Biol. Chem. 270:3429-3434(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91131572; PubMed=1704365;
RA Nies D.E., Hemesath T.J., Kim J.H., Gulcher J.R., Stefansson K.;
RT "The complete cDNA sequence of human hexabrachion (Tenascin). A
RT multidomain protein containing unique epidermal growth factor
RT repeats.";
RL J. Biol. Chem. 266:2818-2823(1991).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92052108; PubMed=1719530;
RA Gulcher J.R., Nies D.E., Alexakos M.J., Ravikant N.A., Sturgill M.E.,
RA Marton L.S., Stefansson K.;
RT "Structure of the human hexabrachion (tenascin) gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:9438-9442(1991).
RN [5]
RP SEQUENCE OF 431-2055 FROM N.A. (ISOFORM 6).
RC TISSUE=Glioblastoma;
RX MEDLINE=89160821; PubMed=2466295;
RA Gulcher J.R., Nies D.E., Marton L.S., Stefansson K.;
RT "An alternatively spliced region of the human hexabrachion contains a
RT repeat of potential N-glycosylation sites.";

```

RL Proc. Natl. Acad. Sci. U.S.A. 86:1588-1592(1989).
RN [6]
RX X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF FIBRONECTIN TYPE-III 3.
RA MSDLIN=93068293; PubMed=1279805;
RA Leahy D.J., Hendrickson W.A., Aukhil I., Erickson H.P.;
RT "Structure of a fibronectin type III domain from tenascin phased by
RT MAD analysis of the selenomethionyl protein.";
RL Science 258:987-991(1992).
CC -1- FUNCTION: SAM (substrate-adhesion molecule) that appears to
CC inhibit cell migration. May play a role in supporting the growth
CC of epithelial tumors. Is a ligand for integrins alpha-8/beta-1,
CC alpha-9/beta-1, alpha-v/beta-3 and alpha-v/beta-6.
CC -1- SUBUNIT: Hexameric. A homotrimer may be formed in the triple
CC coiled-coil region and may be stabilized by disulfide rings at
CC both ends. Two of such half-hexabrachions may be disulfide linked
CC within the central globule.
CC -1- SUBCELLULAR LOCATION: Secreted; extracellular matrix.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=6;
CC Comment=Isoforms are produced in a tissue- and time-specific
CC manner during development;
CC Name=1;
CC IsoId=P24821-1; Sequence=Displayed;
CC Name=2; Synonyms=HT-5;
CC IsoId=P24821-2; Sequence=VSP_001412, VSP_001413;
CC Name=3;
CC IsoId=P24821-3; Sequence=VSP_001412;
CC Name=4; Synonyms=HT-33;
CC IsoId=P24821-4; Sequence=VSP_001413;
CC Name=5;
CC IsoId=P24821-5; Sequence=VSP_001414;
CC Name=6; Synonyms=P31;
CC IsoId=P24821-6; Sequence=VSP_001415;
CC -1- INDUCTION: By TGF-beta.
CC -1- SIMILARITY: Contains 15 EGF-like domains.
CC -1- SIMILARITY: Contains 15 fibronectin type III domains.
CC -1- SIMILARITY: Contains 1 fibrinogen C-terminal domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X56160; CAA39628.1; -
DR EMBL; X78565; CAA53309.1; -
DR EMBL; M55618; AAA88083.1; -
DR EMBL; M24630; AAA52703.1; -
DR PIR; I38337; A32160.
DR PDB; 1TEN; 31-OCT-93.
DR Genew; HGNC:5318; TNC.
DR MIM; 187380; -
DR GO; GO:0005488; F-binding; TAS.
DR GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR GO; GO:0007155; P:cell adhesion; TAS.
DR InterPro; IPR006209; EGF-like.
DR InterPro; IPR002181; Fibrinogen C.
DR InterPro; IPR008957; FN III-like.
DR InterPro; IPR003961; FN-III.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR002049; Laminin_EGF.
DR Pfam; PF00008; EGF; 13.
DR Pfam; PF00147; fibrinogen_C; 1.
DR Pfam; PF00041; fn3; 15.
DR PRINTS; PR00011; EGFLAMININ.
DR SMART; SM00181; EGF; 8.
DR SMART; SM00186; FBS; 1.
DR SMART; SM00060; FN3; 13.
DR PROSITE; PS00022; EGF_1; 15.
DR PROSITE; PS01186; EGF_2; 15.
DR PROSITE; PS50026; EGF_3; 5.

```

```

KW Glycoprotein; Cell adhesion; Repeat; EGF-like domain; Coiled coil;
KW Extracellular matrix; Signal; Alternative splicing; 3D-structure;
FT POLYMORPHISM.
FT CHAIN 1 22
FT CHAIN 23 2201
FT CHAIN 118 145
FT CHAIN 174 186
FT CHAIN 186 217
FT CHAIN 217 248
FT CHAIN 248 280
FT CHAIN 280 311
FT CHAIN 311 342
FT CHAIN 342 373
FT CHAIN 373 404
FT CHAIN 404 435
FT CHAIN 435 466
FT CHAIN 466 497
FT CHAIN 497 528
FT CHAIN 528 559
FT CHAIN 559 590
FT CHAIN 590 621
FT CHAIN 621 710
FT CHAIN 710 801
FT CHAIN 801 891
FT CHAIN 891 983
FT CHAIN 983 1071
FT CHAIN 1071 1162
FT CHAIN 1162 1253
FT CHAIN 1253 1344
FT CHAIN 1344 1435
FT CHAIN 1435 1526
FT CHAIN 1526 1617
FT CHAIN 1617 1708
FT CHAIN 1708 1797
FT CHAIN 1797 1885
FT CHAIN 1885 1973
FT CHAIN 1973 2061
FT CHAIN 2061 2150
FT CHAIN 2150 2240
FT CHAIN 2240 2330
FT CHAIN 2330 2420
FT CHAIN 2420 2510
FT CHAIN 2510 2600
FT CHAIN 2600 2690
FT CHAIN 2690 2780
FT CHAIN 2780 2870
FT CHAIN 2870 2960
FT CHAIN 2960 3050
FT CHAIN 3050 3140
FT CHAIN 3140 3230
FT CHAIN 3230 3320
FT CHAIN 3320 3410
FT CHAIN 3410 3500
FT CHAIN 3500 3590
FT CHAIN 3590 3680
FT CHAIN 3680 3770
FT CHAIN 3770 3860
FT CHAIN 3860 3950
FT CHAIN 3950 4040
FT CHAIN 4040 4130
FT CHAIN 4130 4220
FT CHAIN 4220 4310
FT CHAIN 4310 4400
FT CHAIN 4400 4490
FT CHAIN 4490 4580
FT CHAIN 4580 4670
FT CHAIN 4670 4760
FT CHAIN 4760 4850
FT CHAIN 4850 4940
FT CHAIN 4940 5030
FT CHAIN 5030 5120
FT CHAIN 5120 5210
FT CHAIN 5210 5300
FT CHAIN 5300 5390
FT CHAIN 5390 5480
FT CHAIN 5480 5570
FT CHAIN 5570 5660
FT CHAIN 5660 5750
FT CHAIN 5750 5840
FT CHAIN 5840 5930
FT CHAIN 5930 6020
FT CHAIN 6020 6110
FT CHAIN 6110 6200
FT CHAIN 6200 6290
FT CHAIN 6290 6380
FT CHAIN 6380 6470
FT CHAIN 6470 6560
FT CHAIN 6560 6650
FT CHAIN 6650 6740
FT CHAIN 6740 6830
FT CHAIN 6830 6920
FT CHAIN 6920 7010
FT CHAIN 7010 7100
FT CHAIN 7100 7190
FT CHAIN 7190 7280
FT CHAIN 7280 7370
FT CHAIN 7370 7460
FT CHAIN 7460 7550
FT CHAIN 7550 7640
FT CHAIN 7640 7730
FT CHAIN 7730 7820
FT CHAIN 7820 7910
FT CHAIN 7910 8000
FT CHAIN 8000 8090
FT CHAIN 8090 8180
FT CHAIN 8180 8270
FT CHAIN 8270 8360
FT CHAIN 8360 8450
FT CHAIN 8450 8540
FT CHAIN 8540 8630
FT CHAIN 8630 8720
FT CHAIN 8720 8810
FT CHAIN 8810 8900
FT CHAIN 8900 8990
FT CHAIN 8990 9080
FT CHAIN 9080 9170
FT CHAIN 9170 9260
FT CHAIN 9260 9350
FT CHAIN 9350 9440
FT CHAIN 9440 9530
FT CHAIN 9530 9620
FT CHAIN 9620 9710
FT CHAIN 9710 9800
FT CHAIN 9800 9890
FT CHAIN 9890 9980
FT CHAIN 9980 10070
FT CHAIN 10070 10160
FT CHAIN 10160 10250
FT CHAIN 10250 10340
FT CHAIN 10340 10430
FT CHAIN 10430 10520
FT CHAIN 10520 10610
FT CHAIN 10610 10700
FT CHAIN 10700 10790
FT CHAIN 10790 10880
FT CHAIN 10880 10970
FT CHAIN 10970 11060
FT CHAIN 11060 11150
FT CHAIN 11150 11240
FT CHAIN 11240 11330
FT CHAIN 11330 11420
FT CHAIN 11420 11510
FT CHAIN 11510 11600
FT CHAIN 11600 11690
FT CHAIN 11690 11780
FT CHAIN 11780 11870
FT CHAIN 11870 11960
FT CHAIN 11960 12050
FT CHAIN 12050 12140
FT CHAIN 12140 12230
FT CHAIN 12230 12320
FT CHAIN 12320 12410
FT CHAIN 12410 12500
FT CHAIN 12500 12590
FT CHAIN 12590 12680
FT CHAIN 12680 12770
FT CHAIN 12770 12860
FT CHAIN 12860 12950
FT CHAIN 12950 13040
FT CHAIN 13040 13130
FT CHAIN 13130 13220
FT CHAIN 13220 13310
FT CHAIN 13310 13400
FT CHAIN 13400 13490
FT CHAIN 13490 13580
FT CHAIN 13580 13670
FT CHAIN 13670 13760
FT CHAIN 13760 13850
FT CHAIN 13850 13940
FT CHAIN 13940 14030
FT CHAIN 14030 14120
FT CHAIN 14120 14210
FT CHAIN 14210 14300
FT CHAIN 14300 14390
FT CHAIN 14390 14480
FT CHAIN 14480 14570
FT CHAIN 14570 14660
FT CHAIN 14660 14750
FT CHAIN 14750 14840
FT CHAIN 14840 14930
FT CHAIN 14930 15020
FT CHAIN 15020 15110
FT CHAIN 15110 15200
FT CHAIN 15200 15290
FT CHAIN 15290 15380
FT CHAIN 15380 15470
FT CHAIN 15470 15560
FT CHAIN 15560 15650
FT CHAIN 15650 15740
FT CHAIN 15740 15830
FT CHAIN 15830 15920
FT CHAIN 15920 16010
FT CHAIN 16010 16100
FT CHAIN 16100 16190
FT CHAIN 16190 16280
FT CHAIN 16280 16370
FT CHAIN 16370 16460
FT CHAIN 16460 16550
FT CHAIN 16550 16640
FT CHAIN 16640 16730
FT CHAIN 16730 16820
FT CHAIN 16820 16910
FT CHAIN 16910 17000
FT CHAIN 17000 17090
FT CHAIN 17090 17180
FT CHAIN 17180 17270
FT CHAIN 17270 17360
FT CHAIN 17360 17450
FT CHAIN 17450 17540
FT CHAIN 17540 17630
FT CHAIN 17630 17720
FT CHAIN 17720 17810
FT CHAIN 17810 17900
FT CHAIN 17900 17990
FT CHAIN 17990 18080
FT CHAIN 18080 18170
FT CHAIN 18170 18260
FT CHAIN 18260 18350
FT CHAIN 18350 18440
FT CHAIN 18440 18530
FT CHAIN 18530 18620
FT CHAIN 18620 18710
FT CHAIN 18710 18800
FT CHAIN 18800 18890
FT CHAIN 18890 18980
FT CHAIN 18980 19070
FT CHAIN 19070 19160
FT CHAIN 19160 19250
FT CHAIN 19250 19340
FT CHAIN 19340 19430
FT CHAIN 19430 19520
FT CHAIN 19520 19610
FT CHAIN 19610 19700
FT CHAIN 19700 19790
FT CHAIN 19790 19880
FT CHAIN 19880 19970
FT CHAIN 19970 20060
FT CHAIN 20060 20150
FT CHAIN 20150 20240
FT CHAIN 20240 20330
FT CHAIN 20330 20420
FT CHAIN 20420 20510
FT CHAIN 20510 20600
FT CHAIN 20600 20690
FT CHAIN 20690 20780
FT CHAIN 20780 20870
FT CHAIN 20870 20960
FT CHAIN 20960 21050
FT CHAIN 21050 21140
FT CHAIN 21140 21230
FT CHAIN 21230 21320
FT CHAIN 21320 21410
FT CHAIN 21410 21500
FT CHAIN 21500 21590
FT CHAIN 21590 21680
FT CHAIN 21680 21770
FT CHAIN 21770 21860
FT CHAIN 21860 21950
FT CHAIN 21950 22040
FT CHAIN 22040 22130
FT CHAIN 22130 22220
FT CHAIN 22220 22310
FT CHAIN 22310 22400
FT CHAIN 22400 22490
FT CHAIN 22490 22580
FT CHAIN 22580 22670
FT CHAIN 22670 22760
FT CHAIN 22760 22850
FT CHAIN 22850 22940
FT CHAIN 22940 23030
FT CHAIN 23030 23120
FT CHAIN 23120 23210
FT CHAIN 23210 23300
FT CHAIN 23300 23390
FT CHAIN 23390 23480
FT CHAIN 23480 23570
FT CHAIN 23570 23660
FT CHAIN 23660 23750
FT CHAIN 23750 23840
FT CHAIN 23840 23930
FT CHAIN 23930 24020
FT CHAIN 24020 24110
FT CHAIN 24110 24200
FT CHAIN 24200 24290
FT CHAIN 24290 24380
FT CHAIN 24380 24470
FT CHAIN 24470 24560
FT CHAIN 24560 24650
FT CHAIN 24650 24740
FT CHAIN 24740 24830
FT CHAIN 24830 24920
FT CHAIN 24920 25010
FT CHAIN 25010 25100
FT CHAIN 25100 25190
FT CHAIN 25190 25280
FT CHAIN 25280 25370
FT CHAIN 25370 25460
FT CHAIN 25460 25550
FT CHAIN 25550 25640
FT CHAIN 25640 25730
FT CHAIN 25730 25820
FT CHAIN 25820 25910
FT CHAIN 25910 26000
FT CHAIN 26000 26090
FT CHAIN 26090 26180
FT CHAIN 26180 26270
FT CHAIN 26270 26360
FT CHAIN 26360 26450
FT CHAIN 26450 26540
FT CHAIN 26540 26630
FT CHAIN 26630 26720
FT CHAIN 26720 26810
FT CHAIN 26810 26900
FT CHAIN 26900 26990
FT CHAIN 26990 27080
FT CHAIN 27080 27170
FT CHAIN 27170 27260
FT CHAIN 27260 27350
FT CHAIN 27350 27440
FT CHAIN 27440 27530
FT CHAIN 27530 27620
FT CHAIN 27620 27710
FT CHAIN 27710 27800
FT CHAIN 27800 27890
FT CHAIN 27890 27980
FT CHAIN 27980 28070
FT CHAIN 28070 28160
FT CHAIN 28160 28250
FT CHAIN 28250 28340
FT CHAIN 28340 28430
FT CHAIN 28430 28520
FT CHAIN 28520 28610
FT CHAIN 28610 28700
FT CHAIN 28700 28790
FT CHAIN 28790 28880
FT CHAIN 28880 28970
FT CHAIN 28970 29060
FT CHAIN 29060 29150
FT CHAIN 29150 29240
FT CHAIN 29240 29330
FT CHAIN 29330 29420
FT CHAIN 29420 29510
FT CHAIN 29510 29600
FT CHAIN 29600 29690
FT CHAIN 29690 29780
FT CHAIN 29780 29870
FT CHAIN 29870 29960
FT CHAIN 29960 30050
FT CHAIN 30050 30140
FT CHAIN 30140 30230
FT CHAIN 30230 30320
FT CHAIN 30320 30410
FT CHAIN 30410 30500
FT CHAIN 30500 30590
FT CHAIN 30590 30680
FT CHAIN 30680 30770
FT CHAIN 30770 30860
FT CHAIN 30860 30950
FT CHAIN 30950 31040
FT CHAIN 31040 31130
FT CHAIN 31130 31220
FT CHAIN 31220 31310
FT CHAIN 31310 31400
FT CHAIN 31400 31490
FT CHAIN 31490 31580
FT CHAIN 31580 31670
FT CHAIN 31670 31760
FT CHAIN 31760 31850
FT CHAIN 31850 31940
FT CHAIN 31940 32030
FT CHAIN 32030 32120
FT CHAIN 32120 32210
FT CHAIN 32210 32300
FT CHAIN 32300 32390
FT CHAIN 32390 32480
FT CHAIN 32480 32570
FT CHAIN 32570 32660
FT CHAIN 32660 32750
FT CHAIN 32750 32840
FT CHAIN 32840 32930
FT CHAIN 32930 33020
FT CHAIN 33020 33110
FT CHAIN 33110 33200
FT CHAIN 33200 33290
FT CHAIN 33290 33380
FT CHAIN 33380 33470
FT CHAIN 33470 33560
FT CHAIN 33560 33650
FT CHAIN 33650 33740
FT CHAIN 33740 33830
FT CHAIN 33830 33920
FT CHAIN 33920 34010
FT CHAIN 34010 34100
FT CHAIN 34100 34190
FT CHAIN 34190 34280
FT CHAIN 34280 34370
FT CHAIN 34370 34460
FT CHAIN 34460 34550
FT CHAIN 34550 34640
FT CHAIN 34640 34730
FT CHAIN 34730 34820
FT CHAIN 34820 34910
FT CHAIN 34910 35000
FT CHAIN 35000 35090
FT CHAIN 35090 35180
FT CHAIN 35180 35270
FT CHAIN 35270 35360
FT CHAIN 35360 35450
FT CHAIN 35450 35540
FT CHAIN 35540 35630
FT CHAIN 35630 35720
FT CHAIN 35720 35810
FT CHAIN 35810 35900
FT CHAIN 35900 35990
FT CHAIN 35990 36080
FT CHAIN 36080 36170
FT CHAIN 36170 36260
FT CHAIN 36260 36350
FT CHAIN 36350 36440
FT CHAIN 36440 36530
FT CHAIN 36530 36620
FT CHAIN 36620 36710
FT CHAIN 36710 36800
FT CHAIN 36800 36890
FT CHAIN 36890 36980
FT CHAIN 36980 37070
FT CHAIN 37070 37160
FT CHAIN 37160 37250
FT CHAIN 37250 37340
FT CHAIN 37340 37430
FT CHAIN 37430 37520
FT CHAIN 37520 37610
FT CHAIN 37610 37700
FT CHAIN 37700 37790
FT CHAIN 37790 37880
FT CHAIN 37880 37970
FT CHAIN 37970 38060
FT CHAIN 38060 38150
FT CHAIN 38150 38240
FT CHAIN 38240 38330
FT CHAIN 38330 38420
FT CHAIN 38420 38510
FT CHAIN 38510 38600
FT CHAIN 38600 38690
FT CHAIN 38690 38780
FT CHAIN 38780 38870
FT CHAIN 38870 38960
FT CHAIN 38960 39050
FT CHAIN 39050 39140
FT CHAIN 39140 39230
FT CHAIN 39230 39320
FT CHAIN 39320 39410
FT CHAIN 39410 39500
FT CHAIN 39500 39590
FT CHAIN 39590 39680
FT CHAIN 39680 39770
FT CHAIN 39770 39860
FT CHAIN 39860 39950
FT CHAIN 39950 40040
FT CHAIN 40040 40130
FT CHAIN 40130 40220
FT CHAIN 40220 40310
FT CHAIN 40310 40400
FT CHAIN 40400 40490
FT CHAIN 40490 40580
FT CHAIN 40580 40670
FT CHAIN 40670 40760
FT CHAIN 40760 40850
FT CHAIN 40850 40940
FT CHAIN 40940 41030
FT CHAIN 41030 41120
FT CHAIN 41120 41210
FT CHAIN 41210 41300
FT CHAIN 41300 41390
FT CHAIN 41390 41480
FT CHAIN 41480 41570
FT CHAIN 41570 41660
FT CHAIN 41660 41750
FT CHAIN 41750 41840
FT CHAIN 41840 41930
FT CHAIN 41930 42020
FT CHAIN 42020 42110
FT CHAIN 42110 42200
FT CHAIN 42200 42290
FT CHAIN 42290 42380
FT CHAIN 42380 42470
FT CHAIN 42470 42560
FT CHAIN 42560 42650
FT CHAIN 42650 42740
FT CHAIN 42740 42830
FT CHAIN 42830 42920
FT CHAIN 42920 43010
FT CHAIN 43010 43100
FT CHAIN 43100 43190
FT CHAIN 43190 43280
FT CHAIN 43280 43370
FT CHAIN 43370 43460
FT CHAIN 43460 43550
FT CHAIN 43550 43640
FT CHAIN 43640 43730
FT CHAIN 43730 43820
FT CHAIN 43820 43910
FT CHAIN 43910 44000
FT CHAIN 44000 44090
FT CHAIN 44090 44180
FT CHAIN 44180 44270
FT CHAIN 44270 44360
FT CHAIN 44360 44450
FT CHAIN 44450 44540
FT CHAIN 44540 44630
FT CHAIN 44630 44720
FT CHAIN 44720 44810
FT CHAIN 44810 44900
FT CHAIN 44900 44990
FT CHAIN 44990 45080
FT CHAIN 45080 45170
FT CHAIN 45170 45260
FT CHAIN 45260 45350
FT CHAIN 45350 45440
FT CHAIN 45440 45530
FT CHAIN 45530 45620
FT CHAIN 45620 45710
FT CHAIN 45710 45800
FT CHAIN 45800 45890
FT CHAIN 45890 45980
FT CHAIN 45980 46070
FT CHAIN 46070 46160
FT CHAIN 46160 46250
FT CHAIN 46250 46340
FT CHAIN 46340 46430
FT CHAIN 46430 46520
FT CHAIN 46520 46610
FT CHAIN 46610 46700
FT CHAIN 46700 46790
FT CHAIN 46790 46880
FT CHAIN 46880 46970
FT CHAIN 46970 47060
FT CHAIN 47060 47150
FT CHAIN 47150 47240
FT CHAIN 47240 47330
FT CHAIN 47330 47420
FT CHAIN 47420 47510
FT CHAIN 47510 47600
FT CHAIN 47600 47690
FT CHAIN 47690 47780
FT CHAIN 47780 47870
FT CHAIN 47870 47960
FT CHAIN 47960 48050
FT CHAIN 48050 48140
FT CHAIN 48140 48230
FT CHAIN 48230 48320
FT CHAIN 48320 48410
FT CHAIN 48410 48500
FT CHAIN 48500 48590
FT CHAIN 48590 48680
FT CHAIN 48680 48770
FT CHAIN 48770 48860
FT CHAIN 48860 48950
FT CHAIN 48950 49040
FT CHAIN 49040 49130
FT CHAIN 49130 49220
FT CHAIN 49220 49310
FT CHAIN 49310 49400
FT CHAIN 49400 49490
FT CHAIN 49490 49580
FT CHAIN 49580 49670
FT CHAIN 49670 49760
FT CHAIN 49760 49850
FT CHAIN 49850 49940
FT CHAIN 49940 50030
FT CHAIN 50030 50120
FT CHAIN 50120 50210
FT CHAIN 50210 50300
FT CHAIN 50300 50390
FT CHAIN 50390 50480
FT CHAIN 50480 50570
FT CHAIN 50570 50660
FT CHAIN 50660 50750
FT CHAIN 50750 50840
FT CHAIN 50840 50930
FT CHAIN 50930 51020
FT CHAIN 51020 51110
FT CHAIN 51110 51200
FT CHAIN 51200 51290
FT CHAIN 51290 51380
FT CHAIN 51380 51470
FT CHAIN 51470 51560
FT CHAIN 51560 51650
FT CHAIN 51650 51740
FT CHAIN 51740 51830
FT CHAIN 51830 51920
FT CHAIN 51920 52010
FT CHAIN 52010 52100
FT CHAIN 52100 52190
FT CHAIN 52190 52280
FT CHAIN 52280 52370
FT CHAIN 52370 52460
FT CHAIN 52460 52550
FT CHAIN 52550 52640
FT CHAIN 52640 52730
FT CHAIN 52730 52820
FT CHAIN 52820 52910
FT CHAIN 52910 53000
FT CHAIN 53000 53090
FT CHAIN 53090 53180
FT CHAIN 53180 53270
FT CHAIN 53270 53360
FT CHAIN 53360 53450
FT CHAIN 53450 53540
FT CHAIN 53540 53630
FT CHAIN 53630 53720
FT CHAIN 53720 53810
FT CHAIN 53810 53900
FT CHAIN 53900 53990
FT CHAIN 53990 54080
FT CHAIN 54080 54170
FT CHAIN 54170 54260
FT CHAIN 54260 54350
FT CHAIN 54350 54440
FT CHAIN 54440 54530
FT CHAIN 54530 54620
FT CHAIN 54620 54710
FT CHAIN 54710 54800
FT CHAIN 54800 54890
FT CHAIN 54890 54980
FT CHAIN 54980 55070
FT CHAIN 55070 55160
FT CHAIN 55160 55250
FT CHAIN 55250 55340
FT CHAIN 55340 55430
FT CHAIN 55430 55520
FT CHAIN 55520 55610
FT CHAIN 55610 55700
FT CHAIN 55700 55790
FT CHAIN 55790 55880
FT CHAIN 55880 55970
FT CHAIN 55970 56060
FT CHAIN 56060 56150
FT CHAIN 56150 56240
FT CHAIN 56240 56330
FT CHAIN 56330 56420
FT CHAIN 56420 56510
FT CHAIN 56510 56600
FT CHAIN 56600 56690
FT CHAIN 56690 56780
FT CHAIN 56780 56870
FT CHAIN 56870 56960
FT CHAIN 56960 57050
FT CHAIN 57050 57140
FT CHAIN 57140 57230
FT CHAIN 57230 57320
FT CHAIN 57320 57410
FT CHAIN 57410 57500
FT CHAIN 57500 57590
FT CHAIN 57590 57680
FT CHAIN 57680 57770
FT CHAIN 57770 57860
FT CHAIN 57860 57950
FT CHAIN 57950 58040
FT CHAIN 58040 58130
FT CHAIN 58130 58220
FT CHAIN 58220 58310
FT CHAIN 58310 58400
FT CHAIN 58400 58490
FT CHAIN 58490 58580
FT CHAIN 58580 58670
FT CHAIN 58670 58760
FT CHAIN 58760 58850
FT CHAIN 58850 58940
FT CHAIN 58940 59030
FT CHAIN 59030 59120
FT CHAIN 59120 59210
FT CHAIN 59210 59300
FT CHAIN 59300 59390
FT CHAIN 59390 59480
FT CHAIN 59480 59570
FT CHAIN 59570 59660
FT CHAIN 59660 59750
FT CHAIN 59750 59840
FT CHAIN 59840 59930
FT CHAIN 59930 60020
FT CHAIN 60020 60110
FT CHAIN 60110 60200
FT CHAIN 60200 60290
FT CHAIN 60290 60380
FT CHAIN 60380 60470
FT CHAIN 60470 60560
FT CHAIN 60560 60650
FT CHAIN 60650 60740
FT CHAIN 60740 60830
FT CHAIN 60830 60920
FT CHAIN 60920 61010
FT CHAIN 61010 61100
FT CHAIN 61100 61190
FT CHAIN 61190 61280
FT CHAIN 61280 61370
FT CHAIN 61370 61460
FT CHAIN 61460 
```

FT DISULFID 563 573 BY SIMILARITY.
 FT DISULFID 567 578 BY SIMILARITY.
 Query Match 4.4%; Score 649.5; DB 1; Length 2201;
 Best Local Similarity 22.3%; Pred. No. 3.8e-29;
 Matches 310; Conservative 166; Mismatches 492; Indels 423; Gaps 64;
 QY 545 NDGKSEWSEFLTTAIESVNCPCNVCYNGDCTGTCFLGLHLPDGRASCPVICS-G 603
 Db 208 DDGFTGSDCSQA-----CPSDCNDQKCVNGVCICFEGYAGADCSREICFVPCSEE 259
 QY 604 NGQYMKGRCLCHSGWKAECVPTNOCIDVACSNHGTCTITGTCINPGYKGESCEVDGM 663
 Db 260 HGTGVDGLVCHDFAGDCKPL--CLN-NCYNGRCVNEVCDEGFTGEGDCESELICP 316
 QY 664 DPTCSGRGVVGRSCHFCVGVGGNCTPRATCLDQCSGHGTLPDGLCSDPSPWTGHD 723
 Db 317 N-DCDFRGRCTNGTCYCEGFTGDCGK--TCPHACTGRC--BEGQVCVDEGAGLD 371
 QY 724 CSIEICADGGHCVGCTCRCEDEGMGAACQORACHPRCAEHGTCRDKGKCSQWNG 793
 Db 372 CSEKRCPADCHNRGRVDRGCEDDGFTGADCGELKCPNGCSGHGRVCNQCVCDEGYT 431
 QY 784 EHCT-----IAHYLDKRWKEG-----CPGLCNGNGRCITLDLNGHCV 820
 Db 432 EDCSOLRCPNDCHSRGCV-EGKVCBQCGFKGYDCSDMSCFNDCHQHGRC--VNGM-CV 486
 QY 821 COLGWRGAGC-----DTSMETACGDSKNDGDLV--DCMDPDC-----CL--Q 860
 Db 487 CDDGTGEDCRDRCPRDCNRGLCVDGQCVCEDGFTGPDCAELSCFNDCHGQRCVNGQ 546
 QY 861 PLCH---INPLCLGSPNLDIIQETQVPVSQNLHSEYDRIKFLVGRDSTHIIIPGENPFD 917
 Db 547 CVCHGFMGDKCEQRCPSDCHGQRCVDCQCICHEGF-----TGLDC----- 589
 QY 918 GGHAC-----VIRGQWTSOGTFLVGNVIFSVNPLFGYTISSQDGSFLVYNGGI 968
 Db 590 GQHSQPCDNCNLCQVSGRCICNEG--YSGEDCSEVSPK-----DLV----- 630
 QY 969 SIILRPERAPITQ--EHTLWLPND-RFFVMEITLHREHEIIPSCDLSNFARPNVSP 1025
 Db 631 -----VTEVTSEIWLNDENRVTYELV-----VYTP 658
 QY 1026 SPLTSFASCAEKGFIPEIQALQEEISISGCKMRLSYLSRTPGYSKVLRIISL---THP 1082
 Db 659 T-----HEGLEWQFRVPDQTS-----TIIQELBPGVEYFIRVFAILENKK 700
 QY 1083 TIPNLMKVLHMAVEGELPRKWFAPADLSYFIWDXTV-----YKQVFG- 1130
 Db 701 SIPVSARVATYLPAPGLKPK-----SIKETSVEVEWDPLDAFETWELIIFRNNKDEBGE 756
 QY 1131 -----LSFAVSQVGEVSCPDLI-----LWEKTTVLQG---YEI---D 1164
 Db 757 ITKSLRPEITSYRGTGLAPQOEYHISLHVKNNTRGPKLKRVTITRLDAPSOIEVKDVT 816
 QY 1165 ASKLGGM-SLDKHLANIOSGI-----LHKNGENQFVSQQPPVIGSIMNGRRR 1213
 Db 817 TTALITWFKPLAEIDGIELTYGKIDVPDRTTDLTDEENQY-----SIGNLKP-DTEY 869
 QY 1214 SISCPSCNGLADGNKLAVALTCGSGSLYVGVDFNIRIPSGNVNITLRLNKRDFR 1273
 Db 870 EVSLISRGDMSPN--AKETFTGLDAP-----RNLRRVSQDINSITL-----EWRN 915
 QY 1274 SHSPAHIYLATDPMGAVFLSDSNRRVFKIKSVVVKDLVKNSEVAVAGTQCQLPFD 1333
 Db 916 GKAAIDSYRIKAPISGG-----DHAEDVDFKPSQQA----- 946
 QY 1334 TRCGDGKATEATITNPRGITVDFKGLIYFVGTMTIRIDONGIISTLLGSNDLTARPL 1393
 Db 947 -----TTKTITLIGURPGT--EYGI-----GVSARKEDKENPATINATELDTPKOL 991
 QY 1394 SCDSDMDISQVRLEWPTDLA-----IN-----PMDNSLYLDDNNWVQISE 1434

Db 992 QVSETAETS-LTLLWKTPLAKEFDRYRLNYSLPTGQWGVQVLPFRNTTSVLRG---LEPGQ 1047
 QY 1435 NHQVRIVAGRPMHC-----QVFGIDHFLLSKVAIHATLESATAL--AVSHNGYL 1481
 Db 1048 EYNVLLTAEKGRHKSAPARVKAQTQAPELENLTVEVGWDGLRLNWTAAOQAYEH--- 1103
 QY 1482 YTAETDE-KKINRIQVTTSGEISLVAPAGSCDCNDANCFSGDDGYAKDAKLTPS 1540
 Db 1104 FTIQOEANKVEAARNLTVPG--SURAVDIPLGLKAATPYTVSIYGVIOGY-----RTPV 1155
 QY 1541 SLAVACADGGLYVADLGNIRIRIRKKNKPFNTQNNMYELSSPIDQELYLFDTTGKHLYTQS 1600
 Db 1156 LSAEASTGB--TPNLGEVVVAEVGWDALKLN-----WT 1186
 QY 1601 LPTGDLVLYNFTYTGSDITLITDNNGNMNVRRDSTGMLPLMLVDPGQVYVWMTGNSAL 1660
 Db 1187 APEGAYEYFFIQVQEAADTVEAAQN-----LTPVGLRSTDLPLGLKAAT 1229
 QY 1661 K-SVITQGHSLAMMTYHNGSLATKSNENGWTTTFEYDSFGRLTNVTFPTQGVSSPRSD 1719
 Db 1230 HTYIIRIGVQTDFSTPLSVEVLTEVPDMGNLITVEVSWALRLNWTTPDGTQDFT-- 1287
 QY 1720 TDSVHVQVETSSKODVTITTNLSAGAFYTLQDQVRNSYYIGADGSLRLLLANGMEVA 1779
 Db 1288 -----IQVQEA-----DQVEAHNLTPVGSILR-----SMEI- 1313
 QY 1780 LQTEPHLLAGT 1790
 Db 1314 ----PGLRAGT 1320
 RESULT 3
 ID TENA CHICK STANDARD; PRT; 1808 AA.
 AC F10039; 073584; 073585; P13132;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neuronectin)
 DE (GEM) (J1) (Morfendinous antigen) (Glioma-associated-extracellular
 DE matrix antigen) (GP 150-225).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=90030407; PubMed=2478295;
 RA Spring J., Beck K., Chiquet-Ehrismann R.;
 RT "Two contrary functions of tenascin: dissection of the active sites
 RT by recombinant tenascin fragments.";
 RL Cell 59:325-334(1989).
 RN [2]
 RP SEQUENCE OF 27-722 FROM N.A., AND SEQUENCE OF 79-96.
 RC TISSUE=Fibroblast;
 RX MEDLINE=89030589; PubMed=2460335;
 RA Pearson C.A., Pearson D., Shibahara S., Hofsteenge J.,
 RA Chiquet-Ehrismann R.;
 RT "tenascin: cDNA cloning and induction by TGF-beta.";
 RL EMBO J. 7:2977-2982(1988).
 RN [3]
 RP SEQUENCE OF 464-1018 AND 1412-1661 FROM N.A., AND SEQUENCE OF 852-868.
 RC TISSUE=Embryo;
 RX MEDLINE=88176910; PubMed=2451243;
 RA Jones F.S., Burgoon M.P., Hoffman S., Crossin K.L., Cunningham B.A.,
 RA Edelman G.M.;
 RT "A cDNA clone for cytotactin contains sequences similar to epidermal
 RT growth factor-like repeats and segments of fibronectin and
 RT fibrinogen.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:2186-2190(1988).
 CC -!- FUNCTION: SAM (substrate-adhesion molecule) that appears to

FT CONFLICT 182 182 /FTID=VSP_001411.
FT CONFLICT 563 571 W -> R (IN REF. 2).
FT CONFLICT 598 598 SCNDCCNV -> PAMTATTW (IN REF. 3).
FT CONFLICT 838 838 E -> G (IN REF. 3).
FT CONFLICT 886 886 T -> TEY (IN REF. 3).
FT CONFLICT 886 886 N -> F (IN REF. 3).
SQ SEQUENCE 1808 AA, 198858 MW, B924A06CF9EFD6DE CRC64;

Query Match 4.3%; Score 644; DB 1; Length 1808;
Best Local Similarity 20.5%; Pred. No. 5.7e-29;
Matches 334; Conservative 219; Mismatches 566; Indels 510; Gaps 71;

QY 543 FYNDGSEVSVFLTTAIESVDNCPNVCYNGGDCISGTCCHCFLGFLGPGDGRASCPLVCS 602
DB 332 FCEGYTGDCGELT-----CPNNGNGRCENGLVCVCHGFGVDDDCSQKRCPCDCN 383
QY 603 GNGQYMKGRCLCHSGWKAEC---VPTNQCIDVACSNHGTCTTGTCINPGYKGESCE 659
DB 384 NRGHCVDRGVCHEGLGDCGELRCFND-----CHNRGRGNCGQVDCDEGFIGEDCGE 437
QY 660 VDCNDPTCSRGVYRGECHCFVWGTCNCETPRATCLDOCSGHGFLPDTGLCLCDPSW 719
DB 438 LRCFN-DCHNRGRVNGQCHGFGFEDCGELR--CPNDCNHRGRCV--NGQCVDBEGY 492
QY 720 TGHDCSTIECAADCGHGVGCTCRCEGDMGAACDQACHPRCAEHGTCRDGKCECSP 779
DB 493 TGEDCGELRCFNDCHNRGRVCEGRVCYDNGFMGDCGELSCPNDCQHGRCVDRGCVCHE 552
QY 780 GWNGEHTIAHYLDVVYKBCGPGCLNGRCCTLDLNGWHVCQGLWRGACDTSMETACG 839
DB 553 GFTGEDCR-----ERSCPNDCNVGRC---VEG-RCVCEGYMG-----587
QY 840 DSKNDGDGLVDCMDPCCLQPLCHINPLCLGSPNLDIIQETQVPVSQNLASFYDRI- 898
DB 588 -----IDCSD-----VSP-----PTELVTNVTKTV--NLEWKHENLV 619
QY 899 -KFLV-----GRDSTHIIPE-----NPFDDGHACVIRGOWMTSDGTPLVGNISF 943
DB 620 NEYLTVVPTSSGGLDQFTVPGNOTSATIHELEPGVEYFIRVFAILKXKSI-----672
QY 944 VNNPLFGYTSRQSGFDLVNGGISILIRFERAPFITQSHLWLPWDRF--FVMEII 1000
DB 673 -----PVSARVATVLAPEG-----LKFKS-----VRETSQVQVWDPLSISFDGWELV 715
QY 1001 MRHEENIPSCDL-SNFARNPVVSPSPLTSFASSCAEKGPVPEIOALQEEISISGCKM 1059
DB 716 FRNMOKXDDNGDITSLKRE-----TSYMPQGLAPG-----QYQNVS--L 754
QY 1060 RLSYLSRTPGYKSVLRISLTHPTIPFNLMMKVLHMAVEGRLEFRKWFAPAAPDLSYFIWD 1119
DB 755 HIVKNTRGPGLSRVITTKDAPS-----QIEAK-----D 784
QY 1120 KTDVYNQVFLSEAFVSGVEYESCPLILWEKRTTVLQGE-----IDA 1165
DB 785 VTDT-----TALITWSFLAEIEGIELTYGKDPVGDRTTIDL 822
QY 1166 SKLGWSLDKHALNIOGILHKGNGNQFVSQOFPVIGSMGNRRS---ISCPCNG 1222
DB 823 SE-----DENQY-----SIGNLRPHTEYEVTLISRRG 849
QY 1223 LADGNKLLAPVALTCGSDGLYVGFNVIRIFPSGNVNIILELRNKFDRHSHSPAKEY 1282
DB 850 DMESDPA-----KEVFVTDLPAPRLKXVSQTDNSITL---EWNKSHANIDNYR 895
QY 1283 LATDPMGAVFLSDNSRRVFKIXSTVVVVDLVNKEVWAGTGQCLPFPDTRCGDGGA 1342
DB 896 IKFAPISG-----DHTELTVPKNQA-----T 918
QY 1343 TEALITNPRGTVTKFGLIYFVDGTMIRRIDONGIISTLLGSNLTISARPLSCDVMDIS 1402
DB 919 TRATLTGLRP-----GTEYIGVTAVERQ-DRESAPATINAGTDLDPKDLVSDPTE-T 970
QY 1403 QVRLEWPTDLA-----INPMDNSLYLDNNVNLQISENHQVRIVAG 1443

RESULT 4

TENX HUMAN STANDARD; PRT: 4289 AA.
ID TENX HUMAN P22105; P78531; Q08424; Q9UMG7;
AC P22105; P78531; Q08424; Q9UMG7;
DT 01-AUG-1991 (Rel. 19, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Tenascin X precursor (TN-X) (Hexabrachion-like).
CS TNXB OR TNX OR XB OR HXBL.
CS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_Taxid:9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Rowen L., Danks C., Baskin D., Faust J., Loretz C., Ahearn M.E.,
RA Banta A., Schwartzell S., Smith T.M., Spies T., Hood L.,
RT "Sequence determination of 300 kilobases of the human class III MHC
locus.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]

FT	DOMAIN	2266	2347	FIBROECTIN	TYPE-III 15.
FT	DOMAIN	2365	2446	FIBROECTIN	TYPE-III 16.
FT	DOMAIN	2471	2552	FIBROECTIN	TYPE-III 17.
FT	DOMAIN	2582	2663	FIBROECTIN	TYPE-III 18.
FT	DOMAIN	2688	2769	FIBROECTIN	TYPE-III 19.
FT	DOMAIN	2794	2875	FIBROECTIN	TYPE-III 20.
FT	DOMAIN	2889	2972	FIBROECTIN	TYPE-III 21.
FT	DOMAIN	2997	3078	FIBROECTIN	TYPE-III 22.
FT	DOMAIN	3105	3186	FIBROECTIN	TYPE-III 23.
FT	DOMAIN	3211	3292	FIBROECTIN	TYPE-III 24.
FT	DOMAIN	3307	3384	FIBROECTIN	TYPE-III 25.
FT	DOMAIN	3399	3481	FIBROECTIN	TYPE-III 26.
FT	DOMAIN	3494	3575	FIBROECTIN	TYPE-III 27.
FT	DOMAIN	3601	3682	FIBROECTIN	TYPE-III 28.
FT	DOMAIN	3699	3787	FIBROECTIN	TYPE-III 29.
FT	DOMAIN	3801	3879	FIBROECTIN	TYPE-III 30.
FT	DOMAIN	3890	3971	FIBROECTIN	TYPE-III 31.
FT	DOMAIN	3978	4059	FIBROECTIN	TYPE-III 32.
FT	DOMAIN	4071	4289	FIBROGEN C-TERMINAL.	
FT	SITE	1748	1750	CELL ATTACHMENT SITE (POTENTIAL).	
FT	DISULFID	187	197	BY SIMILARITY.	
FT	DISULFID	191	202	BY SIMILARITY.	
FT	DISULFID	204	213	BY SIMILARITY.	
FT	DISULFID	218	228	BY SIMILARITY.	
FT	DISULFID	222	233	BY SIMILARITY.	
FT	DISULFID	235	244	BY SIMILARITY.	
FT	DISULFID	249	259	BY SIMILARITY.	
FT	DISULFID	253	264	BY SIMILARITY.	
FT	DISULFID	266	275	BY SIMILARITY.	
FT	DISULFID	280	290	BY SIMILARITY.	
FT	DISULFID	284	295	BY SIMILARITY.	
FT	DISULFID	297	306	BY SIMILARITY.	
FT	DISULFID	311	321	BY SIMILARITY.	
FT	DISULFID	315	326	BY SIMILARITY.	
FT	DISULFID	328	337	BY SIMILARITY.	

Query Match 4.0%; Score 597.5; DB 1; Length 4289;
 Best Local Similarity 20.7%; Pred. No. 1.2e-25;
 Matches 325; Conservative 159; Mismatches 515; Indels 571; Gaps 66;

QY	562	SVNCPNVCYNGDCISGTHCHFLGFLGPCGRASCPVLCSGNGQYWKGRCLCHSGWGA	621
DB	369	STRCPDRCDGRGRCDEGICDGTGSGDGCGRVSCPCGNCGRCDRCVCWPGYGT	428
QY	622	ECDVPTNQCIDVACSNHGTCTGTICINPGYKGSCEBVDMDPTCSGRGVCVRGCHCF	681
DB	429	DCG---SRAPRCDRCGRGRCNGVVCNAGYSGEDCGVRS-C-PGDCRGRGRCESGRMCW	484
QY	682	VWGGRNCT-PRATCLD-----	712
DB	485	PGYTRGDCGTRACPGDCRGRGRCVDRGRCVNCNPGTGTGDCGRRCPGDCRCHG--LCEDGV	542
QY	713	CSDPSTWTHDCSIEICAADCGHGVCGTCTCEGDMGAACDQACHPRCAEHGTCRD	772
DB	543	CVCDAVSGBDCSTRSCPGGCRGRGQCLDRCVCEGSDGCGVQCPNDCSQHGVCD	602
QY	773	GKCESPGMNGEHCITIAHYLDVRYKGCPCGLCNGRCRTLDLNGHWCVCQOLGWRGACD	832
DB	603	GVCIWEGYVSDCSI-----RTCPNCHGRGRC---BEGRCLCDPGYGTGTCAT	649
QY	833	SMETACSDKNDGDGLVDCMDPCCLP--LCHN-----	877
DB	650	RMCPA-----DCRGRGRCVQVCLCHVGYGSDCGCBEPFACPGCGPRE	696
QY	878	IIO-----	890
DB	597	LCRAGQCVVEGFGPPCAIQTCEGDCRGRGCHDSCVCKDYGAGDCGEARVPSSA--	754
QY	891	LHSPYDRIFLVRGDSHTIIPGNPPDGGHACVIRGQVMTSDGTPLVGNISFVNPLFG	950
DB	755	--SAYDQ-----RGLAPGE-----YQVTVRALRGTSWGLPASKTITTMIDGP---	795
QY	951	YTIISQDSFDLVINGGISILIRAPERAPFITQHTLWLPW-----DRFFV-----	995

DB	796	QDLRVAVT-----PITTELGWLRPOAEVDVRFVSVVSAGNQ	832
QY	996	-----METII-----MRHEENEIFSCOLSN	1015
DB	833	RVRLEVPPEADGTLTLDLMPGVEYVVTVAERGAVSYPASVRANTEEREESP----	886
QY	1016	FARNPVVSPSLTSPASSCAKEGPIVPEIQALGEEISISCKMRLSVLSRRTGYKSVL	1075
DB	887	---PPSLSQPPRRPWGNTLAELSFRFGTVQDLERHLRAHGYPLR---ANQT---YTSVA	937
QY	1076	RISLTHPTIPENLMKVH-----LMVAVEG-----RLFRKWFATAAP----	1110
DB	938	R-----HIHEYLOQVLGSSADGALLVSLDGLRGQFVRLWRPQPABGP	984
QY	1111	---DLSYFIMDKTDVYNQKVFGLSEAFVSVGYEVESCPDLLILWEKRTVLGCEIDASKL	1168
DB	985	GGELT-----VPGTTRTV-----SLPDL---RPGTT---YHVEVHG	1015
QY	1169	GWMSLDKHALNIOGILHKNGENQFVSQPPVIGSIMGNR-----	1211
DB	1016	RAGQTSKSYAFTTTPSTTQGAQAPLLQRPQBELGELRVLGRDETGLRVVWTAQDPTF	1075
QY	1212	-----RRSISCP-----SCNGLADGNKLLAPVALTCG	1238
DB	1076	AVFQLMRVPEPGAHEEVLPGDVQALVPPPPPGTPLYELSLHGVPVPGKPSDIIYQGI	1135
QY	1239	SD-----GSLYV---GDFNYIRIFPSGNVTN-ILELRNKDFRHSHPA	1278
DB	1136	MDKBEKPGKSGPPRLGELTVTRTSDSLLRVTVPEGEPSFVIQYKDRGGQVQVVPV	1195
QY	1279	HKYILATDPMGSAVFLS--DSNSRRVFKIKSTVWVKDLVKNSEVVAGTGDQCLPDDTRCG	1337
DB	1196	E-----QPORSAVITSLDPCRKYKFLYGVGKK---RHGLVAEA--KILPQSDPSPG	1244
QY	1338	DGGKATEATLNPRGITVDKFLYFY-----DGTMRIRIDQNG-----II	1378
DB	1245	TPPHLGNLWVTD-TPDSLHLSWTVEGQDFTFMVQYRDRDRGPRQVVPVEGPERSEW	1301
QY	1379	S-----TLGSDNLTSARPLSCDSV-----MDI	1401
DB	1302	SLSDPDHXYRFLFGIANKKRYGLTADGTTAPERKEPPRPPBEFLQPLLGELTVGTVP	1361
QY	1402	SOVRLWPTDLAINPMDNSLYLDNNVYLQISE---NHQVRIVAGRMHCQVPGIDHFL	1458
DB	1362	DSLRLSW--TVAQGPF-----SFMVQYKDAQGQVAVAGDENEVTVPGLDPRK	1411
QY	1459	SKVALHATL-----ESATALAVSHNGV-----LYIAETDEKKINRIQVITSG	1501
DB	1412	YKMNLYGLRGQRVGPESVWAKTAPQEDVETPSTELGTAPESPESPL--LGELTVTG	1469
QY	1502	---EISLVAGAPSCDCKDANCDCFSGDDGYAKDAKLNTPSSLAVCADGELYVADLGN	1557
DB	1470	SSPDSLFLWTVPG-----SFDSTVQYKDRGRPRAVRVGKSESVTVGG	1516
QY	1558	IRIRIRKKNKPLNTQNMVE-----LSSPIDQ---ELYLFDT	1591
DB	1517	LE---PGHKYKCHLYGLHEGQRVGFSVAVGTAPQEEETPPATESPLERLGELTVTDV	1572
QY	1592	TGKHI--YTQSLPTGDYLYNFTYTGDDITLITDNG--NMVNVRRDSTGMLWLVV-1645	
DB	1573	TPNSVGLSWTVPEGGF-----DSFVQYKDKQGVQVVAADQREVTVYNLEPERK	1624
QY	1646	---DGQYVWTVMTNSALKSVTTQGHLEAMMTYHNSGLLATKSNENGTWTTYE	1696
DB	1625	YKMNMYGLHDGQ---RMGPLSV--IVTAPATEASKPPLERLGLT-----VTDIT	1671
QY	1697	YDSFGLTNVTFPTQVSSF---RSDTSSVHVQVETSSKDDVTITNLSAGAFYTL--	1752
DB	1672	PDVSG--LSWTVPEGEFDSFVQYKDRDQGVQVVAADQREVTVYNLEPERK	1728
QY	1753	--QDOVRNS 1759	

FT	DOMAIN	221	258	EGF-LIKE 6.	EGF-LIKE 6.	CALCIUM-BINDING (POTENTIAL).	DB	915	CLANPCQNGSGMDGVNTFSCLPLPGFTGDKQCDIMNECLSEPCKNGTCSYVNTCK	974
FT	DOMAIN	260	296	EGF-LIKE 7.	EGF-LIKE 7.	CALCIUM-BINDING (POTENTIAL).	QY	715	CDPSWTGHC--SIEICA-ADCGHGVGVGG---TCRCEDGWMGAAC---DQACHPR	763
FT	DOMAIN	298	336	EGF-LIKE 8.	EGF-LIKE 8.	CALCIUM-BINDING (POTENTIAL).	DB	975	COAGPDGVHCENNINECTESSFCNGTCTVDGINSFSCCLPCVGTGSCFLHEINESSHP-	1033
FT	DOMAIN	375	413	EGF-LIKE 10.	EGF-LIKE 10.	CALCIUM-BINDING (POTENTIAL).	QY	764	CAERGTCTDGG---KCECSPGNGEHC-TIAHYLDVVVKEGCPGLCNGNGCTLDLNGWH	818
FT	DOMAIN	415	454	EGF-LIKE 11.	EGF-LIKE 11.	CALCIUM-BINDING (POTENTIAL).	DB	1034	CLNEGTCTVDGLTGYRCSCPLGYTGKNCOTLVNLSR-----SFCNKGTCTCVOKKAESQ	1086
FT	DOMAIN	456	492	EGF-LIKE 12.	EGF-LIKE 12.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	494	530	EGF-LIKE 13.	EGF-LIKE 13.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	532	568	EGF-LIKE 14.	EGF-LIKE 14.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	570	605	EGF-LIKE 15.	EGF-LIKE 15.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	607	643	EGF-LIKE 16.	EGF-LIKE 16.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	645	680	EGF-LIKE 17.	EGF-LIKE 17.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	682	718	EGF-LIKE 18.	EGF-LIKE 18.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	720	755	EGF-LIKE 19.	EGF-LIKE 19.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	757	793	EGF-LIKE 20.	EGF-LIKE 20.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	795	831	EGF-LIKE 21.	EGF-LIKE 21.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	833	871	EGF-LIKE 22.	EGF-LIKE 22.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	873	909	EGF-LIKE 23.	EGF-LIKE 23.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	911	947	EGF-LIKE 24.	EGF-LIKE 24.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	949	985	EGF-LIKE 25.	EGF-LIKE 25.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	987	1023	EGF-LIKE 26.	EGF-LIKE 26.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1025	1061	EGF-LIKE 27.	EGF-LIKE 27.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1063	1099	EGF-LIKE 28.	EGF-LIKE 28.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1101	1147	EGF-LIKE 29.	EGF-LIKE 29.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1149	1185	EGF-LIKE 30.	EGF-LIKE 30.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1187	1223	EGF-LIKE 31.	EGF-LIKE 31.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1225	1262	EGF-LIKE 32.	EGF-LIKE 32.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1264	1302	EGF-LIKE 33.	EGF-LIKE 33.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1304	1343	EGF-LIKE 34.	EGF-LIKE 34.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1374	1412	EGF-LIKE 35.	EGF-LIKE 35.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1420	1456	LIN/NOTCH 1.	LIN/NOTCH 1.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1456	1492	LIN/NOTCH 2.	LIN/NOTCH 2.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1503	1535	ANK 1.	ANK 1.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1535	1571	ANK 2.	ANK 2.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1571	1607	ANK 3.	ANK 3.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1607	1643	ANK 4.	ANK 4.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1643	1679	ANK 5.	ANK 5.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1679	1715	ANK 6.	ANK 6.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1715	1751	POLY-ALA.	POLY-ALA.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1751	1787	POLY-LEU.	POLY-LEU.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1787	1823	POLY-SER.	POLY-SER.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1823	1859	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1859	1895	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1895	1931	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	1931	1967	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	1967	2003	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	2003	2039	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	2039	2075	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	2075	2111	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	2111	2147	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	2147	2183	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	2183	2219	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	2219	2255	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	2255	2291	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	2291	2327	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	2327	2363	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	2363	2399	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	2399	2435	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	2435	2471	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	2471	2507	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131
FT	DOMAIN	2507	2543	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	QY	819	CVCQLGMRGAGCDTSMETACGDSKNDGDGLVDCWDPCCLQPLCHINPLCLGSPN	874
FT	DOMAIN	2543	2579	BY SIMILARITY.	BY SIMILARITY.	CALCIUM-BINDING (POTENTIAL).	DB	1087	CLCPSGAGACVDP-NVSCDIAASRRG-----VLVEHLQHSVGVCINAGN	1131

Query Match 3.0%; Score 440; DB 1; Length 2471;

Best Local Similarity 30.1%; Pred. No. 6.8e-17; Matches 107; Conservative 44; Mismatches 143; Indels 62; Gaps 20;

QY 562 SVNCPSPN-CYNGDC---ISG-TCHCFLGFLGPDGCRASCP-----VLCSNGGOY 607

DB 795 NIDECASNPCLNQTCTPDISGVTCHCVLPYTGKNCQTVLAPCPSPCNAAVCKESP 854

QY 608 MKRCICLCHSWKAGCDVFNQCIDVACSNHFC--ITGT--CICNPGVKGSCFP--VD 661

DB 855 ESYTCLCAPGQGCRTCTIDICISKPCNMHGLCHNTQSGMYCECPFGSGMDCEDIDD 914

QY 662 CNDPTCSRGVCRG-----EHCFCVGVGNGTNCETPRATCLDQ-CSGHGTF--LPDTGLCS 714

RESULT 6

NOTC DROME

ID NOTC DROME STANDARD; PRT; 2703 AA.

AC P07207; 097458; P04154; Q9W4T8;

DT 01-NOV-1986 (Rel. 03, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Neurogenic locus Notch protein precursor.

GN N OR EG:140G11.1 OR EG:163A10.2 OR CG3936.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_Taxid=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Oregon-R; TISSUE=Embryo;

RX MEDLINE=86079539; PubMed=3935325;

RA Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S.;

RT "Nucleotide sequence from the neurogenic locus notch implies a gene

RT product that shares homology with proteins containing EGF-like

RT repeats.";

RL Cell 43:567-581(1985).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Canton-S, and Oregon-R; TISSUE=Embryo;

RX MEDLINE=87064624; PubMed=3097517;

RA Kidd S., Kelley M.R., Young M.W.;

RT "Sequence of the notch locus of Drosophila melanogaster: relationship

RT of the encoded protein to mammalian clotting and growth factors.";

RL Mol. Cell. Biol. 6:3094-3108(1986).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA April J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,

RA Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busan D.A., Butler H., Cadenot E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Danke C., Davidson L.B., Davies P.,

RA De Fabios B., Deicher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,

RA Hoshin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sider-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Spradling M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*";
RT Science 287:2185-2195(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R;
RX MEDLINE=20196011; PubMed=10731137;
RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
RA Barrell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,
RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert P., Borkova D.,
RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
RA Papagiannakis G., Spanos L., Cox S., Madueno E., de Pablo B.,
RA Modrell J., Peter A., Schoettler P., Werner M., Mourikoti F.,
RA Beinert N., Dowe G., Schaefer U., Jaeckle H., Bucheton A.,
RA Callister D.M., Campbell L.A., Darlamitsou A., Henderson N.S.,
RA McMillan P.J., Sallies C., Tait E.A., Valenti P., Saunders R.D.C.,
RA Glover D.M.;
RT "From sequence to chromosome: the tip of the X chromosome of *D. melanogaster*";
RT Science 287:2220-2222(2000).
RN [5]
RP SEQUENCE OF 2505-2611 FROM N.A.
RX MEDLINE=85093329; PubMed=2981631;
RA Wharton K.A., Yedvobnick B., Finnerty V.G., Artavanis-Tsakonas S.;
RT "opa: a novel family of transcribed repeats shared by the Notch locus and other developmentally regulated loci in *D. melanogaster*";
RL Cell 40:55-62(1985).
RN [6]
RP SEQUENCE OF 1-8 FROM N.A.
RX MEDLINE=87257846; PubMed=3037327;
RA Kelley M.R., Kidd S., Berg R.L., Young M.W.;
RT "Restriction of P-element insertions at the Notch locus of *Drosophila melanogaster*";
RL Mol. Cell. Biol. 7:1545-1548(1987).
RN [7]
RP INTERACTION WITH DX, AND MUTANT SU42C.
RX MEDLINE=94215489; PubMed=8162848;
RA Diederich R.J., Matsuno K., Hing H., Artavanis-Tsakonas S.;
RT "Cytosolic interaction between *delex* and Notch ankyrin repeats implicates *delex* in the Notch signaling pathway";
RL Development 120:473-481(1994).
RN [8]
RP INTERACTION WITH DX.
RX MEDLINE=95401878; PubMed=7671825;
RA Matsuno K., Diederich R.J., Go M.J., Blaumueller C.M.,
RA Artavanis-Tsakonas S.;
RT "Deltex acts as a positive regulator of Notch signaling through interactions with the Notch ankyrin repeats";
RL Development 121:2633-2644(1995).
RN [9]
RP S3 CLEAVAGE BY PSN.
RX MEDLINE=99221487; PubMed=10206646;
RA Struhl G., Greenwald I.,
RT "Presenilin is required for activity and nuclear access of Notch in *Drosophila*";
RL Nature 398:522-525(1999).
RN [10]
RP S3 CLEAVAGE BY PSN.
RX MEDLINE=99221488; PubMed=10206647;
RA Ye Y., Lukinova N., Fortini M.E.;

RT "Neurogenic phenotypes and altered Notch processing in *Drosophila* Presenilin mutants";
RL Nature 398:525-529(1999).
RN [11]
RP S2 CLEAVAGE BY KUZ.
RX MEDLINE=21657146; PubMed=11799064;
RA Lieber T., Kidd S., Young M.W.;
RT "kuzbanian-mediated cleavage of *Drosophila* Notch";
RL Genes Dev. 16:209-221(2002).
RN [12]
RP MUTANT MCD5.
RX MEDLINE=21575956; PubMed=11719214;
RA Raman P., Knechtman K., Seugnet L., Arbogast N., Ackermann C.,
RA Heitzler P.;
RT "Novel Notch alleles reveal a Deltex-dependent pathway repressing neural fate";
RT Curr. Biol. 11:1729-1738(2001).
RN [13]
RP REVIEW
RX MEDLINE=22256570; PubMed=12369105;
RA Portin P.;
RT "General outlines of the molecular genetics of the Notch signalling pathway in *Drosophila melanogaster*: a review";
RL Hereditas 136:89-96(2002).
CC -!- FUNCTION: Signaling protein, which regulates, with both positive and negative signals, the differentiation of at least central and peripheral nervous system and eye, wing disk, oogenesis, segmental appendages such as antennae and legs, and muscles, through lateral inhibition or induction. Functions as a receptor for membrane-bound ligands Delta and Serrate to regulate cell-fate determination. Upon ligand activation, and releasing from the cell membrane, the Notch intracellular domain (NICD) forms a transcriptional activator complex with Su(H) (Suppressor of hairless) and activates genes of the *E(spl)* complex. Essential for proper differentiation of ectoderm.
CC -!- SUBUNIT: Interacts with Su(H) when activated. Interacts with Dx via its ANK repeats.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Upon activation and S3 cleavage, it is released from the cell membrane and enters into the nucleus in conjunction with Su(H).
CC -!- PTM: Upon binding its ligands such as Delta or Serrate, it is cleaved (S2 cleavage) in its extracellular domain, close to the transmembrane domain. S2 cleavage is probably mediated by Kuz. It is then cleaved (S3 cleavage) downstream of its transmembrane domain, releasing it from the cell membrane. S3 cleavage requires Pen.
CC -!- SIMILARITY: Belongs to the NOTCH family.
CC -!- SIMILARITY: Contains 36 EGF-like domains.
CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.
CC -!- SIMILARITY: Contains 6 ANK repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M16152; AAB59220.1; -
CC EMBL; M16153; AAB59220.1; JOINED.
CC EMBL; M16149; AAB59220.1; JOINED.
CC EMBL; M16150; AAB59220.1; JOINED.
CC EMBL; M16151; AAB59220.1; JOINED.
CC EMBL; K03508; AAA28725.1; -
CC EMBL; M13689; AAA28725.1; JOINED.
CC EMBL; K03507; AAA28725.1; JOINED.
CC EMBL; AE003426; AAF45848.2; -
CC EMBL; AL035436; CAB37610.1; -
CC EMBL; M12175; AAA74496.1; -
CC EMBL; M16025; AAA28726.1; -

Query Match		2.9%; Score 438.5; DB 1; Length 2703;
Best Local Similarity		29.7%; Pred. No. 9.6e-17;
Matches 125; Conservative 50; Mismatches 165; Indels 81; Gaps 24;		
QY	561	ESVDNCP--SNYGVNGDCI-----SGTCHQFLGFLGPDGCR-----ASCPLVSGNGQYMK 609
DB	448	EDIDEDQSGPCEHNGICVNTPSYRCNSQSGTGPRCTNINECHSHP--CQNEGSCLD 505
QY	610	G-----RLCHSGWKGAECVDFTNQCIDVACSNHGTC---ITG-TCTCNFGYKGSCE---E 659
DB	506	DPGTFRCVCMFPGTGTQCEIDIDECQSNPCLNDGTCHDKINGPKSCALGFTGARQINI 565
QY	660	VDCWMDPTCSGRGVC---VRG-ECHCFVGMGTGNCETPRATCLDQCSGHGTFELPDITG--LC 713
DB	566	DDCQSPFCNRNGICHDSIAGYSCCEPPGYGTGTSCININDCNSNPCHRGKCIDDVNSFKC 625
QY	714	SCDPSMTGHDGDCIEI--CAAD-CGGHGVC---YGG-TCRCEDGMMGAACDQRA--CHPR- 763
DB	626	LCDFGVTGYICQIQCINESNPFCDFHCDQDRGVSYCQAGTSGKNCENVNNECHSNP 685
QY	764	CAHGTCRDG---KCECSFGWNGEHTAHYLDRLVVKGCPCPLGNGRCGLDNLNGWHC 819
DB	686	CNNGATCIDGINSYKQCVGFTGQHC-----EKNVDECISSPCANNVGVCIDQVNGYKC 739
QY	820	VCQLGWRGAGCDTSMETACGDSKNDG---DGLVDQMDPDCCLQPLCHINPLCLGSPNPL 876
DB	740	ECRFGFYDAHLSDVDEACSNPCVNEGRCEGNEF-----ICHCFPGYTKRCCL 790
QY	877	DIQETQVPSVQNLHSGFYRIKFLVGRDSTHIIPG-----ENPFDGGHAC 922
DB	791	DIDECSSNPC--GHGGTCYDKLNAF-----SCQCMPGYTGKQKCTNIDDCVTNPGCNGGTC 844
QY	923	V 923
DB	845	I 845
RESULT 7		
ID	JAG3 BRARE	STANDARD; PRT; 1213 AA.
AC	Q30Y54;	2003 (Rel. 41, Created)
DT	28-FEB-2003	(Rel. 41, Last sequence update)
DT	10-OCT-2003	(Rel. 42, Last annotation update)
DE	Jagged 3 precursor	(Jagged3).
GN	JAG3.	
OS	Brachydanio rerio (Zebrafish)	(Danio rerio).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;	
OC	Cyprinidae; Danio.	
OX	NCBI_TaxID=7955;	
RN	[1]	SEQUENCE FROM N.A.
RP	Oda T., Chandrasekharappa S.C.;	
RA	"Isolation, characterization and expression analysis of Zebrafish	
RT	Jagged genes."	
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.	
CC	-!- FUNCTION: Ligand for Notch receptors and involved in the mediation	
CC	of Notch signaling (By similarity). Seems to be involved in cell-	
CC	fate decisions	
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).	
CC	-!- SIMILARITY: Contains 17 EGF-like domains.	
CC	-!- SIMILARITY: Contains 1 DSL domain.	
CC	-!- SIMILARITY: Contains 1 VWF domain.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC	or send an email to license@isb-sib.ch).	
CC	-----	

DR	EMBL; AF229451; AAL08216.1; --	
DR	ZFIN; ZDB-GENE-011128-4; Jag3.	
DR	InterPro; IPR000152; Asx_Hydroxyl_S.	
DR	InterPro; IPR001774; DSL.	
DR	InterPro; IPR000742; EGF 2.	
DR	InterPro; IPR001881; EGF_Ca.	
DR	InterPro; IPR001438; EGF_II.	
DR	InterPro; IPR006209; EGF_like	
DR	InterPro; IPR002049; Laminin_EGF.	
DR	InterPro; IPR009041; PMP_inhibitor.	
DR	InterPro; IPR001007; VWF_C.	
DR	Pfam; PF01414; DSL; 1.	
DR	Pfam; PF00008; EGF; 14.	
DR	PRINTS; PR00010; EGFBL00D.	
DR	PRINTS; PR00011; EGFAMININ.	
DR	SMART; SM00051; DSL; 1.	
DR	SMART; SM00179; EGF_CA; 10.	
DR	SMART; SM00214; VWC; 1.	
DR	PROSITE; PS00010; ASX_HYDROXYL; 10.	
DR	PROSITE; PS00022; EGF_1; 16.	
DR	PROSITE; PS01186; EGF_2; 12.	
DR	PROSITE; PS00026; EGF_3; 15.	
DR	PROSITE; PS01187; EGF_CA; 8.	
DR	PROSITE; PS01208; VWF_C; 1; FALSE NEG.	
KW	Calcium-binding; EGF-like domain; Glycoprotein; Developmental protein;	
KW	Repeat; Transmembrane; Signal.	
FT	SIGNAL	1 26
FT	CHAIN	27 1213
FT	DOMAIN	27 1064
FT	TRANSMEM	1065 1087
FT	DOMAIN	1088 1213
FT	DOMAIN	154 226
FT	DOMAIN	227 260
FT	DOMAIN	258 291
FT	DOMAIN	293 331
FT	DOMAIN	333 369
FT	DOMAIN	371 407
FT	DOMAIN	409 445
FT	DOMAIN	447 482
FT	DOMAIN	484 520
FT	DOMAIN	522 558
FT	DOMAIN	552 624
FT	DOMAIN	626 662
FT	DOMAIN	664 700
FT	DOMAIN	702 738
FT	DOMAIN	746 777
FT	DOMAIN	779 815
FT	DOMAIN	817 853
FT	DOMAIN	860 914
FT	DOMAIN	918 956
FT	DISULFID	231 242
FT	DISULFID	235 248
FT	DISULFID	250 259
FT	DISULFID	262 273
FT	DISULFID	288 279
FT	DISULFID	281 290
FT	DISULFID	297 309
FT	DISULFID	303 319
FT	DISULFID	321 330
FT	DISULFID	337 348
FT	DISULFID	342 357
FT	DISULFID	359 368
FT	DISULFID	375 386
FT	DISULFID	380 395
FT	DISULFID	397 406
FT	DISULFID	413 424
FT	DISULFID	418 433
FT	DISULFID	435 444
FT	DISULFID	451 461
FT	DISULFID	455 470
FT	DISULFID	472 481
FT	DISULFID	488 499
FT	DISULFID	493 508

```
FT DISULFID 510 519 BY SIMILARITY.
FT DISULFID 526 537 BY SIMILARITY.
FT DISULFID 531 546 BY SIMILARITY.
FT DISULFID 548 557 BY SIMILARITY.
FT DISULFID 630 641 BY SIMILARITY.
FT DISULFID 635 650 BY SIMILARITY.
FT DISULFID 652 661 BY SIMILARITY.
FT DISULFID 668 679 BY SIMILARITY.
FT DISULFID 673 688 BY SIMILARITY.
FT DISULFID 690 699 BY SIMILARITY.
FT DISULFID 706 717 BY SIMILARITY.
FT DISULFID 711 726 BY SIMILARITY.
FT DISULFID 728 737 BY SIMILARITY.
FT DISULFID 783 794 BY SIMILARITY.
FT DISULFID 788 803 BY SIMILARITY.
FT DISULFID 805 814 BY SIMILARITY.
FT DISULFID 821 832 BY SIMILARITY.
FT DISULFID 826 841 BY SIMILARITY.
FT DISULFID 843 852 BY SIMILARITY.
FT DOMAIN 938 941 POLY-PRO.
FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 556 556 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 742 742 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 957 957 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 988 988 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1042 1042 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1213 AA; 13365 MW; 565F16A7E20D9534 CRC64;

Query Match 2.9%; Score 434; DB 1; Length 1213;
Best Local Similarity 27.2%; Pred. No. 4.8e-17;
Matches 134; Conservative 48; Mismatches 177; Indels 134; Gaps 27;

QY 496 LLDGRLLLTQEARSLCTPQSGTVP-----PSSHETGFIQYLD-----SG----- 537
Db 98 VLGNKFTSGTRS-----EKSRIVLPFFPAWPSRY-TLIVEALDFNNETASGKLIRK 151
QY 538 IWLAFYNDGKE-----SEVVSF-----LTTAIESV--DNCPSNC-----YGNQDC-I 577
Db 152 AYHSGMINPNQRORLTHNGFVAQFEYQIRVTCLEHYVFGCNKFCRPRDEFHYTCDO 211
QY 578 SGTCHEFLGLPDGRASCPVLCS-----GNGQYMKG----- 610
Db 212 NGNKTCLGWTGPDNTAICRQGCSTEHGSCQKQPGCKLYGQWQGYCDKCIHPHPCVHG 271
QY 611 -----RCLCHGWKAECDDVPNNQC-IDVACSNHGTCT-TG-----TCICNPGYKGESCE 658
Db 272 TCVEPWQCLDTNWGQLCDKDLNYGTHQPCPLNGGTCTNTGPDKVKQCSCEPDYSGVNCE 331
QY 659 EVD--CMDPTCSGRGVVVRG-----EHCFFVGMGGTNCETPRATCL-DQCSHGTF--LPD 709
Db 332 RAEHACLNSPNCANGTCKETSQGYECHCAIGWSGTSCINVDCTPNQCKHGGTCQDLVN 391
QY 710 TGLCSCDPSWTHDCSIE-----ICAADCGHGVCGVGTCTCEDGWMGAACDQRA- 759
Db 392 GFKACAPPHWTGTCQLDANECEKPCVNAKSC--HNLIGAYFCECLPGWSQNCQDININ 449
QY 760 -CHPRCAEHGTCRD----GKCECSPGWNGEHCTIAHYLDRVVYKGGCPGLCNGRCRTLLD 814
Db 450 DCKGQCLNGGTCKDLVNGYRCLCPPGYTGEC-----EKDVDECASSFCLNGRGQDQEV 503
QY 815 NGWHVCVQLGWRGAGGDTSMETACGSKNDG-----DGLVDCMD-----PD 856
Db 504 NGFQCLCPAGFSQQLCQLDID-YCKENPCQNGAQCNFLASDYFCRCPDDYEGKNCSHLKD 562
QY 857 CCLQLPLCHINPLC 869
Db 563 HCTTSCQVIDSC 575

RESULT 8
NOTC_XENLA
ID NOTC_XENLA STANDARD; PRT; 2524 AA.
```

```
AC P21783;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurogenic locus notch protein homolog precursor (NOTCH protein).
GN NOTCH.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=90385285; PubMed=2402639;
RA Coffman C., Harris W., Kintner C.;
RT "Notch, the Xenopus homolog of Drosophila notch.";
RL Science 249:1438-1441(1990).
RN [2]
RP REVISIONS TO 1759-1782.
RA Kintner C.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DEVELOPMENTAL STAGE: Expressed almost uniformly in early embryos.
CC -!- SIMILARITY: Belongs to the NOTCH family.
CC -!- SIMILARITY: Contains 36 EGF-like domains.
CC -!- SIMILARITY: Contains 3 Lin/Notch repeats.
CC -!- SIMILARITY: Contains 6 ANK repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M33874; AAB02039.1; --
DR HSSP; P00740; 1EDM.
DR InterPro; IPSC02110; ANK.
DR InterPro; IPSC00152; Asx_hydroxyl_S.
DR InterPro; IPSC00742; EGF_2.
DR InterPro; IPSC01881; EGF_Ca.
DR InterPro; IPSC01438; EGF_II.
DR InterPro; IPSC06209; EGF_Like.
DR InterPro; IPSC02049; Laminin_EGF.
DR InterPro; IPSC08297; Notch_dom.
DR InterPro; IPSC00800; Notch_dom.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00008; EGF; 36.
DR Pfam; PF00066; notch; 3.
DR PIRSF; PIRSF002279; Notch; 1.
DR PRINTS; PR00010; EGFBL00D.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRINTS; PR01452; NOTCH.
DR SMART; SM00248; ANK; 6.
DR SMART; SM00179; EGF_CA; 24.
DR SMART; SM00004; NL; 2.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 23.
DR PROSITE; PS00022; EGF_1; 34.
DR PROSITE; PS01186; EGF_2; 29.
DR PROSITE; PS50026; EGF_3; 36.
DR PROSITE; PS01187; EGF_CA; 21.
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Signal; Glycoprotein.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 2524 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG.
FT DOMAIN 20 1728 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1729 1750 POTENTIAL.
FT DOMAIN 1751 2524 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 20 57 EGF-LIKE 1.
FT DOMAIN 58 99 EGF-LIKE 2.
```


FT	DOMAIN	102	EGF-LIKE 3.	FT	DISULFID	415	428	BY SIMILARITY.
FT	DOMAIN	141	EGF-LIKE 4.	FT	DISULFID	422	437	BY SIMILARITY.
FT	DOMAIN	179	EGF-LIKE 5.	FT	DISULFID	439	448	BY SIMILARITY.
FT	DOMAIN	217	EGF-LIKE 6.	FT	DISULFID	455	466	BY SIMILARITY.
FT	DOMAIN	256	EGF-LIKE 7.	FT	DISULFID	460	475	BY SIMILARITY.
FT	DOMAIN	292	EGF-LIKE 8.	FT	DISULFID	477	486	BY SIMILARITY.
FT	DOMAIN	332	EGF-LIKE 9.	FT	DISULFID	493	504	BY SIMILARITY.
FT	DOMAIN	370	EGF-LIKE 10.	FT	DISULFID	498	513	BY SIMILARITY.
FT	DOMAIN	409	EGF-LIKE 11.	FT	DISULFID	515	524	BY SIMILARITY.
FT	DOMAIN	449	EGF-LIKE 12.	FT	DISULFID	531	542	BY SIMILARITY.
FT	DOMAIN	487	EGF-LIKE 13.	FT	DISULFID	536	551	BY SIMILARITY.
FT	DOMAIN	525	EGF-LIKE 14.	FT	DISULFID	553	562	BY SIMILARITY.
FT	DOMAIN	563	EGF-LIKE 15.	FT	DISULFID	569	578	BY SIMILARITY.
FT	DOMAIN	600	EGF-LIKE 16.	FT	DISULFID	574	588	BY SIMILARITY.
FT	DOMAIN	638	EGF-LIKE 17.	FT	DISULFID	590	599	BY SIMILARITY.
FT	DOMAIN	675	EGF-LIKE 18.	FT	DISULFID	606	617	BY SIMILARITY.
FT	DOMAIN	713	EGF-LIKE 19.	FT	DISULFID	611	626	BY SIMILARITY.
FT	DOMAIN	751	EGF-LIKE 20.	FT	DISULFID	628	637	BY SIMILARITY.
FT	DOMAIN	788	EGF-LIKE 21.	FT	DISULFID	644	654	BY SIMILARITY.
FT	DOMAIN	826	EGF-LIKE 22.	FT	DISULFID	649	663	BY SIMILARITY.
FT	DOMAIN	866	EGF-LIKE 23.	FT	DISULFID	665	674	BY SIMILARITY.
FT	DOMAIN	904	EGF-LIKE 24.	FT	DISULFID	681	692	BY SIMILARITY.
FT	DOMAIN	942	EGF-LIKE 25.	FT	DISULFID	686	701	BY SIMILARITY.
FT	DOMAIN	980	EGF-LIKE 26.	FT	DISULFID	703	712	BY SIMILARITY.
FT	DOMAIN	1018	EGF-LIKE 27.	FT	DISULFID	719	729	BY SIMILARITY.
FT	DOMAIN	1056	EGF-LIKE 28.	FT	DISULFID	724	738	BY SIMILARITY.
FT	DOMAIN	1094	EGF-LIKE 29.	FT	DISULFID	740	749	BY SIMILARITY.
FT	DOMAIN	1142	EGF-LIKE 30.	FT	DISULFID	756	767	BY SIMILARITY.
FT	DOMAIN	1180	EGF-LIKE 31.	FT	DISULFID	761	776	BY SIMILARITY.
FT	DOMAIN	1218	EGF-LIKE 32.	FT	DISULFID	778	787	BY SIMILARITY.
FT	DOMAIN	1264	EGF-LIKE 33.	FT	DISULFID	794	805	BY SIMILARITY.
FT	DOMAIN	1304	EGF-LIKE 34.	FT	DISULFID	799	814	BY SIMILARITY.
FT	DOMAIN	1346	EGF-LIKE 35.	FT	DISULFID	816	825	BY SIMILARITY.
FT	DOMAIN	1383	EGF-LIKE 36.	FT	DISULFID	832	843	BY SIMILARITY.
FT	DOMAIN	1424	EGF-LIKE 37.	FT	DISULFID	837	854	BY SIMILARITY.
FT	REPEAT	1441	LINE/NOTCH 1.	FT	DISULFID	856	865	BY SIMILARITY.
FT	REPEAT	1479	LINE/NOTCH 2.	FT	DISULFID	872	883	BY SIMILARITY.
FT	REPEAT	1521	LINE/NOTCH 3.	FT	DISULFID	877	892	BY SIMILARITY.
FT	REPEAT	1876	ANK 1.	FT	DISULFID	894	903	BY SIMILARITY.
FT	REPEAT	1924	ANK 2.	FT	DISULFID	910	921	BY SIMILARITY.
FT	REPEAT	1957	ANK 3.	FT	DISULFID	915	930	BY SIMILARITY.
FT	REPEAT	1991	ANK 4.	FT	DISULFID	932	941	BY SIMILARITY.
FT	REPEAT	2024	ANK 5.	FT	DISULFID	986	997	BY SIMILARITY.
FT	REPEAT	2053	ANK 6.	FT	DISULFID	991	1006	BY SIMILARITY.
FT	DISULFID	22	BY SIMILARITY.	FT	DISULFID	1008	1017	BY SIMILARITY.
FT	DISULFID	29	BY SIMILARITY.	FT	DISULFID	1024	1035	BY SIMILARITY.
FT	DISULFID	47	BY SIMILARITY.	FT	DISULFID	1029	1044	BY SIMILARITY.
FT	DISULFID	62	BY SIMILARITY.	FT	DISULFID	1046	1055	BY SIMILARITY.
FT	DISULFID	68	BY SIMILARITY.	FT	DISULFID	1062	1073	BY SIMILARITY.
FT	DISULFID	89	BY SIMILARITY.	FT	DISULFID	1067	1082	BY SIMILARITY.
FT	DISULFID	106	BY SIMILARITY.	FT	DISULFID	1084	1093	BY SIMILARITY.
FT	DISULFID	111	BY SIMILARITY.	FT	DISULFID	1100	1121	BY SIMILARITY.
FT	DISULFID	130	BY SIMILARITY.	FT	DISULFID	1115	1130	BY SIMILARITY.
FT	DISULFID	145	BY SIMILARITY.					
FT	DISULFID	156	BY SIMILARITY.					
FT	DISULFID	167	BY SIMILARITY.					
FT	DISULFID	183	BY SIMILARITY.					
FT	DISULFID	188	BY SIMILARITY.					
FT	DISULFID	205	BY SIMILARITY.					
FT	DISULFID	221	BY SIMILARITY.					
FT	DISULFID	226	BY SIMILARITY.					
FT	DISULFID	244	BY SIMILARITY.					
FT	DISULFID	260	BY SIMILARITY.					
FT	DISULFID	265	BY SIMILARITY.					
FT	DISULFID	282	BY SIMILARITY.					
FT	DISULFID	298	BY SIMILARITY.					
FT	DISULFID	305	BY SIMILARITY.					
FT	DISULFID	322	BY SIMILARITY.					
FT	DISULFID	338	BY SIMILARITY.					
FT	DISULFID	343	BY SIMILARITY.					
FT	DISULFID	360	BY SIMILARITY.					
FT	DISULFID	375	BY SIMILARITY.					
FT	DISULFID	380	BY SIMILARITY.					
FT	DISULFID	399	BY SIMILARITY.					

Query Match 2.9%; Score 429; DB 1; Length 2524;
Best Local Similarity 30.6%; Pred. No. 3.1e-16;
Matches 112; Conservative 38; Mismatches 144; Indels 72; Gaps 21;

QY	562	SVNCPN-CYGNDCISG----	TCHFLGFLGPDGR----	ASCPVLCNGNGQYMKG-	610
DB	489	NIDECASNPCLHNGKIDKINEFRCDCTGFSNLCQHDDECTSTP--	CKNGAKCLDGP	546	
QY	611	---RCLCHSGWGAECVPTTQCIDVACNSHGTCITG----	TCICNPYKGESCEE--	VD	661
DB	547	NSYTCQCTEGFTGRHCEQDINECI	PDPC--HYGCKDGIATFTCLCRPGVTGRCLCNDINE	605	
QY	662	QMDPTCSGRGVCVRGE----	CHCFVWGWTNCET-----	PRATCLDQCSGHGFL	707
DB	606	CLSKPCLNGGQCTDRENGVICTCPKGTGTVNCETKIDDCASNLCNDNGKCIDKIDYE---	662		
QY	708	PDTGLCSDPSPSWTGHDCSIEICAAAD----	CGHGVCV----	GGTCRCEDGWMGAAC--	DQR 758
DB	663	-----CTCEPGYTGKLCNININECDNPNCRNGGTCKDQINGFTCVCPDGHDMCLSEVN	717		

759 ACHPRCAEHGTCRDG-----KCECPGNGBHECTIAHYLDVRWKEGCPGLCNGNGRCTLDL 814
 718 ECNSNPCTHGAACHGVNGYKDCBAGSGSNCIDNN-----NECESNPMCGGTCK-DW 770
 815 NG-WHCVCQGLWRGAGDTSVETACGSKNDGDLVDCMDPDC-CLQP-----LCH-INP 867
 771 TGAYICCKAGFGSPCOTINEGSSNPCLNHGTCIDDVAGYKCNCLPVTGAICSAVLA 830
 868 LCLGSP 873
 831 PCAGSP 836

RESULT 9
 NTC1 HUMAN
 ID NTC1 HUMAN STANDARD; PRT; 2556 AA.
 AC P46531;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neurogenic locus notch homolog protein 1 precursor (Notch 1) (hN1)
 DE (translocation-associated notch protein TAN-1).
 GN NOTCH1 OR TAN1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Mann R.S., Blaumueller C.M., Zagouras P.;
 RT "Complete human notch 1 (hN1) cDNA sequence";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 [2]
 RN SEQUENCE OF 1-2444 FROM N.A.
 RC MEDLINE=91347367; PubMed=1831692;
 RA Ellisen L.W., Bird J., West D.C., Soreng A.L., Reynolds T.C.,
 RA Smith S.D., Sklar J.;
 RT "TAN-1, the human homolog of the Drosophila notch gene, is broken by
 RL chromosomal translocations in T lymphoblastic neoplasms.";
 RL Cell 66:649-661(1991).
 [3]
 RN IDENTIFICATION OF LIGANDS.
 RC MEDLINE=99180765; PubMed=10079256;
 RA Gray G.E., Mann R.S., Mitsiades E., Henrique D., Carcangiu M.-L.,
 RA Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;
 RT "Human ligands of the Notch receptor.";
 RL Am. J. Pathol. 154:785-794(1999).
 [4]
 RN INTERACTION WITH DTX1.
 RC MEDLINE=98250176; PubMed=9590294;
 RA Matsuno K., Eastman D., Mitsiades T., Quinn A.M., Carcangiu M.L.,
 RA Orntlich P., Kadesch T., Artavanis-Tsakonas S.;
 RT "Human deltex is a conserved regulator of Notch signalling.";
 RL Nat. Genet. 19:74-78(1998).
 CC -1- FUNCTION: Functions as a receptor for membrane-bound ligands
 CC Jagged1, Jagged2 and Deltal to regulate cell-fate determination.
 CC Upon ligand activation through the released notch intracellular
 CC domain (NICD) it forms a transcriptional activator complex with
 CC RBP-J kappa and activates genes of the enhancer of split locus.
 CC Affects the implementation of differentiation, proliferation and
 CC apoptotic programs. May be important for normal lymphocyte
 CC function. In altered form, may contribute to transformation or
 CC progression in some T-cell neoplasms. Involved in the maturation
 CC of both CD4+ and CD8+ cells in the thymus (By similarity).
 CC -1- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
 CC terminal fragment N(FC) which are probably linked by disulfide
 CC bonds (By similarity). Interacts with DTX1 and DTX2.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
 CC proteolytical processing NICD is translocated to the nucleus (By
 CC similarity).
 CC -1- TISSUE SPECIFICITY: In fetal tissues most abundant in spleen,

brain stem and lung. Also present in most adult tissues where it
 is found mainly in lymphoid tissues.
 CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
 CC which is proteolytically cleaved by a furin-like convertase in the
 CC trans-Golgi network before it reaches the plasma membrane to yield
 CC an active, ligand-accessible form. Cleavage results in a C-
 CC terminal fragment N(TM) and a N-terminal fragment N(FC). Following
 CC ligand binding, it is cleaved by TNF-alpha converting enzyme
 CC (TACE) to yield a membrane-associated intermediate fragment called
 CC notch extracellular truncation (NEXT). This fragment is then
 CC cleaved by presenilin dependent gamma-secretase to release a
 CC notch-derived peptide containing the intracellular domain (NICD)
 CC from the membrane (By similarity).
 CC -1- PTM: Phosphorylated (By similarity).
 CC -1- DISEASE: NOTCH1 truncation is associated with T-cell acute
 CC lymphoblastic leukemia.
 CC -1- SIMILARITY: Belongs to the NOTCH family.
 CC -1- SIMILARITY: Contains 36 EGF-like domains.
 CC -1- SIMILARITY: Contains 3 Lin/Notch repeats.
 CC -1- SIMILARITY: Contains 5 ANK repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF308602; AAG33848.1; -.
 CC EMBL; M73980; AAA60614.1; -.
 CC HSSP; P00740; 1EDM.
 CC Genew; HGNC:7881; NOTCH1.
 CC MIM; 190199; -.
 CC GO; GO:0016021; C:integral to membrane; NAS.
 CC GO; GO:0003793; F:defense/immunity protein activity; NAS.
 CC GO; GO:0006955; P:immune response; NAS.
 CC InterPro; IPR002110; ANK.
 CC InterPro; IPR000152; Asx_hydroxyl_S.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR001881; EGF_Ca.
 CC InterPro; IPR001438; EGF_II.
 CC InterPro; IPR006209; EGF_Like.
 CC InterPro; IPR002049; Laminin_EGF.
 CC InterPro; IPR008297; Notch.
 CC InterPro; IPR000800; Notch_dom.
 CC Pfam; PF00023; ank; 6.
 CC Pfam; PF00008; EGF; 35.
 CC Pfam; PIRSF00066; notch; 3.
 CC PIRSF; PIRSF002279; Notch; 1.
 CC PRINTS; PR00010; EGFBLLOOD.
 CC PRINTS; PR00011; EGFFLAMININ.
 CC PRINTS; PR01452; NOTCH.
 CC SMART; SM00248; ANK; 6.
 CC SMART; SM00179; EGF_CA; 23.
 CC SMART; SM00004; NL; 3.
 CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
 CC PROSITE; PS50086; ANK_REPEAT; 4.
 CC PROSITE; PS00010; ASX_HYDROXYL; 20.
 CC PROSITE; PS00022; EGF_1; 34.
 CC PROSITE; PS01186; EGF_2; 26.
 CC PROSITE; PS00025; EGF_3; 36.
 CC PROSITE; PS01187; EGF_CA; 18.
 CC Receptor; Transcription regulation; Activator; Differentiation;
 CC Developmental protein; Repeat; ANK repeat; EGF-like domain;
 CC Transmembrane; Glycoprotein; Signal; Phosphorylation.
 CC SIGNAL 1 18 POTENTIAL.
 CC CHAIN 19 2556 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 1.
 CC CHAIN 1722 2556 NOTCH EXTRACELLULAR TRUNCATION (BY
 CC SIMILARITY).
 CC CHAIN 1755 2556 NOTCH INTRACELLULAR DOMAIN (BY
 CC SIMILARITY).
 CC DOMAIN 19 1736 EXTRACELLULAR (POTENTIAL).

FT	TRANSMEM	1737	1757	POTENTIAL.		Matches 111; Conservative 42; Mismatches 143; Indels 70; Gaps 20;
FT	DOMAIN	1758	2556	CYTOPLASMIC (POTENTIAL).		
FT	EGF-LIKE 1.	20	58			
FT	DOMAIN	59	99			
FT	EGF-LIKE 2.	102	139			
FT	DOMAIN	140	176			
FT	EGF-LIKE 3.	178	216			
FT	DOMAIN	218	255			
FT	EGF-LIKE 4.	257	293			
FT	DOMAIN	295	333			
FT	EGF-LIKE 5.	335	371			
FT	DOMAIN	372	410			
FT	EGF-LIKE 6.	412	450			
FT	DOMAIN	452	488			
FT	EGF-LIKE 7.	490	526			
FT	DOMAIN	528	564			
FT	EGF-LIKE 8.	566	601			
FT	DOMAIN	603	639			
FT	EGF-LIKE 9.	641	676			
FT	DOMAIN	678	714			
FT	EGF-LIKE 10.	716	751			
FT	DOMAIN	753	789			
FT	EGF-LIKE 11.	791	827			
FT	DOMAIN	829	868			
FT	EGF-LIKE 12.	870	906			
FT	DOMAIN	908	944			
FT	EGF-LIKE 13.	946	982			
FT	DOMAIN	984	1020			
FT	EGF-LIKE 14.	1022	1058			
FT	DOMAIN	1060	1096			
FT	EGF-LIKE 15.	1098	1144			
FT	DOMAIN	1146	1182			
FT	EGF-LIKE 16.	1184	1220			
FT	DOMAIN	1222	1266			
FT	EGF-LIKE 17.	1268	1306			
FT	DOMAIN	1308	1347			
FT	EGF-LIKE 18.	1349	1385			
FT	DOMAIN	1388	1427			
FT	EGF-LIKE 19.	1446	1481			
FT	DOMAIN	1482	1523			
FT	EGF-LIKE 20.	1524	1563			
FT	DOMAIN	1562	1601			
FT	EGF-LIKE 21.	1601	1641			
FT	DOMAIN	1641	1681			
FT	EGF-LIKE 22.	1681	1721			
FT	DOMAIN	1721	1761			
FT	EGF-LIKE 23.	1761	1801			
FT	DOMAIN	1801	1841			
FT	EGF-LIKE 24.	1841	1881			
FT	DOMAIN	1881	1921			
FT	EGF-LIKE 25.	1921	1961			
FT	DOMAIN	1961	2001			
FT	EGF-LIKE 26.	2001	2041			
FT	DOMAIN	2041	2081			
FT	EGF-LIKE 27.	2081	2121			
FT	DOMAIN	2121	2161			
FT	EGF-LIKE 28.	2161	2201			
FT	DOMAIN	2201	2241			
FT	EGF-LIKE 29.	2241	2281			
FT	DOMAIN	2281	2321			
FT	EGF-LIKE 30.	2321	2361			
FT	DOMAIN	2361	2401			
FT	EGF-LIKE 31.	2401	2441			
FT	DOMAIN	2441	2481			
FT	EGF-LIKE 32.	2481	2521			
FT	DOMAIN	2521	2561			
FT	EGF-LIKE 33.	2561	2601			
FT	DOMAIN	2601	2641			
FT	EGF-LIKE 34.	2641	2681			
FT	DOMAIN	2681	2721			
FT	EGF-LIKE 35.	2721	2761			
FT	DOMAIN	2761	2801			
FT	EGF-LIKE 36.	2801	2841			
FT	DOMAIN	2841	2881			
FT	EGF-LIKE 37.	2881	2921			
FT	DOMAIN	2921	2961			
FT	EGF-LIKE 38.	2961	3001			
FT	DOMAIN	3001	3041			
FT	EGF-LIKE 39.	3041	3081			
FT	DOMAIN	3081	3121			
FT	EGF-LIKE 40.	3121	3161			
FT	DOMAIN	3161	3201			
FT	EGF-LIKE 41.	3201	3241			
FT	DOMAIN	3241	3281			
FT	EGF-LIKE 42.	3281	3321			
FT	DOMAIN	3321	3361			
FT	EGF-LIKE 43.	3361	3401			
FT	DOMAIN	3401	3441			
FT	EGF-LIKE 44.	3441	3481			
FT	DOMAIN	3481	3521			
FT	EGF-LIKE 45.	3521	3561			
FT	DOMAIN	3561	3601			
FT	EGF-LIKE 46.	3601	3641			
FT	DOMAIN	3641	3681			
FT	EGF-LIKE 47.	3681	3721			
FT	DOMAIN	3721	3761			
FT	EGF-LIKE 48.	3761	3801			
FT	DOMAIN	3801	3841			
FT	EGF-LIKE 49.	3841	3881			
FT	DOMAIN	3881	3921			
FT	EGF-LIKE 50.	3921	3961			
FT	DOMAIN	3961	4001			
FT	EGF-LIKE 51.	4001	4041			
FT	DOMAIN	4041	4081			
FT	EGF-LIKE 52.	4081	4121			
FT	DOMAIN	4121	4161			
FT	EGF-LIKE 53.	4161	4201			
FT	DOMAIN	4201	4241			
FT	EGF-LIKE 54.	4241	4281			
FT	DOMAIN	4281	4321			
FT	EGF-LIKE 55.	4321	4361			
FT	DOMAIN	4361	4401			
FT	EGF-LIKE 56.	4401	4441			
FT	DOMAIN	4441	4481			
FT	EGF-LIKE 57.	4481	4521			
FT	DOMAIN	4521	4561			
FT	EGF-LIKE 58.	4561	4601			
FT	DOMAIN	4601	4641			
FT	EGF-LIKE 59.	4641	4681			
FT	DOMAIN	4681	4721			
FT	EGF-LIKE 60.	4721	4761			
FT	DOMAIN	4761	4801			
FT	EGF-LIKE 61.	4801	4841			
FT	DOMAIN	4841	4881			
FT	EGF-LIKE 62.	4881	4921			
FT	DOMAIN	4921	4961			
FT	EGF-LIKE 63.	4961	5001			
FT	DOMAIN	5001	5041			
FT	EGF-LIKE 64.	5041	5081			
FT	DOMAIN	5081	5121			
FT	EGF-LIKE 65.	5121	5161			
FT	DOMAIN	5161	5201			
FT	EGF-LIKE 66.	5201	5241			
FT	DOMAIN	5241	5281			
FT	EGF-LIKE 67.	5281	5321			
FT	DOMAIN	5321	5361			
FT	EGF-LIKE 68.	5361	5401			
FT	DOMAIN	5401	5441			
FT	EGF-LIKE 69.	5441	5481			
FT	DOMAIN	5481	5521			
FT	EGF-LIKE 70.	5521	5561			
FT	DOMAIN	5561	5601			
FT	EGF-LIKE 71.	5601	5641			
FT	DOMAIN	5641	5681			
FT	EGF-LIKE 72.	5681	5721			
FT	DOMAIN	5721	5761			
FT	EGF-LIKE 73.	5761	5801			
FT	DOMAIN	5801	5841			
FT	EGF-LIKE 74.	5841	5881			
FT	DOMAIN	5881	5921			
FT	EGF-LIKE 75.	5921	5961			
FT	DOMAIN	5961	6001			
FT	EGF-LIKE 76.	6001	6041			
FT	DOMAIN	6041	6081			
FT	EGF-LIKE 77.	6081	6121			
FT	DOMAIN	6121	6161			
FT	EGF-LIKE 78.	6161	6201			
FT	DOMAIN	6201	6241			
FT	EGF-LIKE 79.	6241	6281			
FT	DOMAIN	6281	6321			
FT	EGF-LIKE 80.	6321	6361			
FT	DOMAIN	6361	6401			
FT	EGF-LIKE 81.	6401	6441			
FT	DOMAIN	6441	6481			
FT	EGF-LIKE 82.	6481	6521			
FT	DOMAIN	6521	6561			
FT	EGF-LIKE 83.	6561	6601			
FT	DOMAIN	6601	6641			
FT	EGF-LIKE 84.	6641	6681			
FT	DOMAIN	6681	6721			
FT	EGF-LIKE 85.	6721	6761			
FT	DOMAIN	6761	6801			
FT	EGF-LIKE 86.	6801	6841			
FT	DOMAIN	6841	6881			
FT	EGF-LIKE 87.	6881	6921			
FT	DOMAIN	6921	6961			
FT	EGF-LIKE 88.	6961	7001			
FT	DOMAIN	7001	7041			
FT	EGF-LIKE 89.	7041	7081			
FT	DOMAIN	7081	7121			
FT	EGF-LIKE 90.	7121	7161			
FT	DOMAIN	7161	7201			
FT	EGF-LIKE 91.	7201	7241			
FT	DOMAIN	7241	7281			
FT	EGF-LIKE 92.	7281	7321			
FT	DOMAIN	7321	7361			
FT	EGF-LIKE 93.	7361	7401			
FT	DOMAIN	7401	7441			
FT	EGF-LIKE 94.	7441	7481			
FT	DOMAIN	7481	7521			
FT	EGF-LIKE 95.	7521	7561			
FT	DOMAIN	7561	7601			
FT	EGF-LIKE 96.	7601	7641			
FT	DOMAIN	7641	7681			
FT	EGF-LIKE 97.	7681	7721			
FT	DOMAIN	7721	7761			
FT	EGF-LIKE 98.	7761	7801			
FT	DOMAIN	7801	7841			
FT	EGF-LIKE 99.	7841	7881			
FT	DOMAIN	7881	7921			
FT	EGF-LIKE 100.	7921	7961			
FT	DOMAIN	7961	8001			
FT	EGF-LIKE 101.	8001	8041			
FT	DOMAIN	8041	8081			
FT	EGF-LIKE 102.	8081	8121			
FT	DOMAIN	8121	8161			
FT	EGF-LIKE 103.	8161	8201			
FT	DOMAIN	8201	8241			
FT	EGF-LIKE 104.	8241	8281			
FT	DOMAIN	8281	8321			
FT	EGF-LIKE 105.	8321	8361			
FT	DOMAIN	8361	8401			
FT	EGF-LIKE 106.	8401	8441			
FT	DOMAIN	8441	8481			
FT	EGF-LIKE 107.	8481	8521			
FT	DOMAIN	8521	8561			
FT	EGF-LIKE 108.	8561	8601			
FT	DOMAIN	8601	8641			
FT	EGF-LIKE 109.	8641	8681			
FT	DOMAIN	8681	8721			
FT	EGF-LIKE 110.	8721	8761			
FT	DOMAIN	8761	8801			
FT	EGF-LIKE 111.	8801	8841			
FT	DOMAIN	8841	8881			
FT	EGF-LIKE 112.	8881	8921			
FT	DOMAIN	8921	8961			
FT	EGF-LIKE 113.	8961	9001			
FT	DOMAIN	9001	9041			
FT	EGF-LIKE 114.	9041	9081			
FT	DOMAIN	9081	9121			
FT	EGF-LIKE 115.	9121	9161			
FT	DOMAIN	9161	9201			
FT	EGF-LIKE 116.	9201	9241			
FT	DOMAIN	9241	9281			
FT	EGF-LIKE 117.	9281	9321			
FT	DOMAIN	9321	9361			
FT	EGF-LIKE 118.	9361	9401			
FT	DOMAIN	9401	9441			
FT	EGF-LIKE 119.	9441	9481			
FT	DOMAIN	9481	9521			
FT	EGF-LIKE 120.	9521	9561			
FT	DOMAIN	9561	9601			
FT	EGF-LIKE 121.	9601	9641			
FT	DOMAIN	9641	9681			
FT	EGF-LIKE 122.	9681	9721			
FT	DOMAIN	9721	9761			
FT	EGF-LIKE 123.	9761	9801			
FT	DOMAIN	9801	9841			
FT	EGF-LIKE 124.	9841	9881			
FT	DOMAIN					

RC STRAIN=C57BL/6 X CBA; TISSUE=Embryo;
RX MEDLINE=93178563; PubMed=8440332;
RA Lardelli M., Lendahl U.;
RT "Notch A and Notch B-two mouse Notch homologues coexpressed in a
RL wide variety of tissues.";
RN Exp. Cell Res. 204:364-372(1993).
[5]
RN SEQUENCE OF 1659-1673 FROM N.A.
RX MEDLINE=9364499; PubMed=10437788;
RA Lee J.S., Ishimoto A., Yanagawa S.I.;
RT "Murine leukemia provirus-mediated activation of the Notch1 gene leads
RL to induction of HES-1 in a mouse T lymphoma cell line, DL-3.";
RN FEBS Lett. 455:276-280(1999).
[6]
RN SEQUENCE OF 1950-2201 FROM N.A.
RX MEDLINE=9802946; PubMed=9384671;
RA Messerle M., Folio M., Nehls M., Eggert H., Boehm T.;
RT "Dynamic changes in gene expression during in vitro differentiation of
RL mouse embryonic stem cells.";
RN Cytokines Cell. Mol. Ther. 1:139-143(1995).
[7]
RN SEQUENCE OF 1655-1659, CLEAVAGE BY FURIN-LIKE CONVERTASE, AND
RX MUTAGENESIS OF 1651-ARG--ARG-1654.
RA Logeat F., Bessia C., Brou C., LeBail O., Jarriault S., Seidah N.G.,
RT Israel A.;
RL "The Notch1 receptor is cleaved constitutively by a furin-like
RN convertase.";
RN Proc. Natl. Acad. Sci. U.S.A. 95:8108-8112(1998).
[8]
RN PARTIAL SEQUENCE, AND POST-TRANSLATIONAL PROCESSING.
RX MEDLINE=21523956; PubMed=11518718;
RA Saxena M.T., Schroeter E.H., Mumm J.S., Kopan R.;
RT "Murine notch homologs (NL-4) undergo presenilin-dependent
RL proteolysis.";
RN J. Biol. Chem. 276:40268-40273(2001).
[9]
RN POST-TRANSLATIONAL PROCESSING.
RX MEDLINE=21374376; PubMed=11459941;
RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Konjo T.;
RT "Conservation of the biochemical mechanisms of signal transduction
RL among mammalian Notch family members.";
RN Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).
[10]
RN INTERACTION WITH DTX1 AND DTX2.
RX MEDLINE=21123790; PubMed=11226752;
RA Kishi N., Tang Z., Maeda Y., Hirai A., Mo R., Ito M., Suzuki S.,
RA Nakao K., Kinoshita T., Kadesch I., Hul C.-C., Artavanis-Tsakonas S.,
RA Okano H., Matsuno K.;
RT "Murine homologs of deltex define a novel gene family involved in
RL vertebrate Notch signaling and neurogenesis.";
RN Int. J. Dev. Neurosci. 19:21-35(2001).
CC -1- FUNCTION: Functions as a receptor for membrane-bound ligands
CC Jagged1, Jagged2 and Deltal to regulate cell-fate determination.
CC Upon ligand activation through the released notch intracellular
CC domain (NICD) it forms a transcriptional activator complex with
CC RBP-J kappa and activates genes of the enhancer of split locus.
CC Affects the implementation of differentiation, proliferation and
CC apoptotic programs (By similarity). May play an essential role in
CC postimplantation development, probably in some aspect of cell
CC specification and/or differentiation. May be involved in mesoderm
CC development, somite formation and neurogenesis. Involved in the
CC maturation of both CD4+ and CD8+ cells in the thymus.
CC -1- SUBUNIT: Heterodimer of a C-terminal fragment N(TW) and a N-
CC terminal fragment N(EC) which are probably linked by disulfide
CC bonds. Interacts with DTX1 and DTX2.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Following
CC proteolytical processing NICD is translocated to the nucleus.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=Q01705-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q01705-2; Sequence=VSP_001402, VSP_001403, VSP_001404;
CC Note=No experimental confirmation available;
CC -1- TISSUE SPECIFICITY: Highly expressed in the brain, lung and
CC thymus. Expressed at lower levels in the spleen, bone-marrow,
CC spinal cord, eyes, mammary gland, liver, intestine, skeletal
CC muscle, kidney and heart.
CC -1- DEVELOPMENTAL STAGE: First detected in the mesoderm at 7.5 dpc By
CC 8.5 dpc highly expressed in presomitic mesoderm, mesenchyme and
CC endothelial cells, while much lower levels are seen in the
CC neuroepithelium. Between 9.5-10.5 dpc expressed at high levels in
CC the neuroepithelium. At 13.5 dpc expressed in the surface
CC ectoderm eye and developing whisker follicles.
CC -1- PTM: Synthesized in the endoplasmic reticulum as an inactive form
CC which is proteolytically cleaved by a furin-like convertase in the
CC trans-Golgi network before it reaches the plasma membrane to yield
CC an active, ligand-accessible form. Cleavage results in a C-
CC terminal fragment N(TW) and a N-terminal fragment N(EC). Following
CC ligand binding, it is cleaved by TMP-alpha converting enzyme called
CC (TACE) to yield a membrane-associated intermediate fragment called
CC notch extracellular truncation (NEXT). This fragment is then
CC cleaved by presenilin dependent gamma-secretase to release a
CC notch-derived peptide containing the intracellular domain (NICD)
CC from the membrane.
CC -1- PTM: Phosphorylated.
CC -1- SIMILARITY: Belongs to the NOTCH family.
CC -1- SIMILARITY: Contains 36 EGF-like domains.
CC -1- SIMILARITY: Contains 3 Lin/Notch repeats.
CC -1- SIMILARITY: Contains 5 ANK repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sb-sib.ch).
CC
CC -----
CC EMBL; Z11886; CAA77941.1; -
CC EMBL; L02613; AAK14898.1; -
CC EMBL; X68278; CAA45339.1; -
CC EMBL; AJ238029; CAB40733.1; -
CC EMBL; X82562; CAA57909.1; -
CC PIR; A46019; A46019.
CC PIR; B49175; B49175.
CC HSSP; P00740; 1EDM.
CC MGI; MGI:97363; Notch1.
CC GO; GO:0005887; C:integral to plasma membrane; IC.
CC GO; GO:0005515; F:protein binding; IPI.
CC GO; GO:0030154; P:cell differentiation; IMP.
CC GO; GO:0007386; P:compartment specification; IMP.
CC GO; GO:0007219; P:N signaling pathway; IC.
CC GO; GO:0045944; P:positive regulation of transcription from P. .; IDA.
CC InterPro; IPR002110; ANK.
CC InterPro; IPR00152; Asx_hydroxyl_S.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR001438; EGF_II.
CC InterPro; IPR006209; EGF-like.
CC InterPro; IPR002049; Laminin_EGF.
CC InterPro; IPR008297; Notch_dom.
CC InterPro; IPR008000; Notch_dom.
CC Pfam; PF00023; ank; 7.
CC Pfam; PF00008; EGF; 35.
CC Pfam; PF00065; notch; 3.
CC PIRSF; PIRSF002279; Notch; 1.
CC PRINTS; PR00010; EGFBCLOD.
CC PRINTS; PR00011; EGFFLAMININ.
CC PRINTS; PR01452; NOTCH.
CC SMART; SM00248; ANK; 6.
CC SMART; SM00179; EGF_CA; 24.
CC SMART; SM00004; NL; 2.
CC PROSITE; PS50297; ANK_REPEAT_REGION; 1.
CC PROSITE; PS50089; ANK_REPEAT; 2.

```

DR PROSITE; PS00010; ASX_HYDROXYL; 22.
DR PROSITE; PS00022; EGF_1; 34.
DR PROSITE; PS01186; EGF_2; 27.
DR PROSITE; PS00026; EGF_3; 36.
DR PROSITE; PS01187; EGF_CA; 21.
KW Receptor; Transcription regulation; Activator; Differentiation;
KW Developmental protein; Repeat; ANK repeat; EGF-like domain;
KW Transmembrane; Glycoprotein; Signal; Phosphorylation;
KW Alternative splicing.
FT SIGNAL 1 18
FT CHAIN 19 2531
FT CHAIN 1711 2531
FT CHAIN 1744 2531
FT DOMAIN 19 1725

Query Match 2.9%; Score 427.5; DB 1; Length 2531;
Best Local Similarity 30.2%; Pred. No. 3.8e-16;
Matches 108; Conservative 42; Mismatches 137; Indels 71; Gaps 19;

QY 562 SVNCPSPN-CVGNDCISG-----TCHCFLGLFLGPDQCR-----ASCPVLCSGNGQYWKG- 610
Db 490 NTDECASSPCLHNGHCKMKIHEFOCQCPKGFNGHLCOYDVECASTP--CKNGAKCLDGP 547
QY 611 ---RCLCHSGWKAECDPVPTNQCIDVACSNHGTCITG---TICINPGYKGESCEE--VD 661
Db 548 NTYTCVCTEGVTGTHCEVIDEDCDPDC-HYGSCKDGVATFTCLQCPGYTGHHCEITNINE 606
QY 662 CMPTCSGRGVCRGE-----CHCFVGMGGTNCB-----TPRATCLDQCSHGFTL 707
Db 607 CHSQPCRHGTGTCQDRNSYLCCLKGTTGPNCEINLDDCASNFCDSGTCLDKIDGYE--- 663
QY 708 PDTGLCSDPSWTGHDCSIEI--CA-ADCGGHGVC---VGGTCRCEDGWMGAAC--DQR 758
Db 664 -----CACRGYTGSMNVNIDECAGSPCHNGTCEGDIAGFTCRPEGHDTCLSEVN 718
QY 759 ACHPRCAEHGTCRDG-----KCECSPGWNGEHCTIAHYLDRVKVGCPGLCNGNCRCTLDL 814
Db 719 ECNSNPICIHGACRDGLNGYKCDCAPGWSGTNCDDNN-----NECSNPVWGVTCKDWT 772
QY 815 NGWHVCVQLGWRGAGDCTSMETACGDSKNDGDLVDCMDPDC---CLQPLCHINPLC 869
Db 773 SGIVCTCREGFSGPNQCTNINECASNPCLNQG-----TCDDVAGYKNCNCPPLPYTGATC 826

RESULT 11
NTC2 MOUSE
ID NTC2_MOUSE STANDARD; PRT; 2470 AA.
AC O35516; Q06008; Q60941;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurogenic locus notch homolog protein 2 precursor (Notch 2) (Notch
DE B).
GN NOTCH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Thymus;
RA Hamada Y., Higuchi M., Tsujimoto Y.;
RT "Complete amino acid sequence and multiform transcripts encoded by a
RT single copy of mouse Notch2 gene."
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 316-1518 FROM N.A.
RC STRAIN=C57BL/6 X CBA; TISSUE=Embryo;
RA Lardelli M., Lendahl U.;
RT "Notch A and Notch B-two mouse Notch homologues coexpressed in a
RT wide variety of tissues."
RL Exp. Cell Res. 204:364-372(1993).

```

```

RN [3]
RP SEQUENCE OF 1765-2153 FROM N.A.
RX MEDLINE=97075110; PubMed=8917536;
RA Milner L.A., Bigas A., Kopan R., Brashem-Stein C., Bernstein I.D.,
RA Martin D.I.,
RT "Inhibition of granulocytic differentiation by mNotch1."
RL Proc. Natl. Acad. Sci. U.S.A. 93:13014-13019(1996).
RN [4]
RP FUNCTION.
RX MEDLINE=99396706; PubMed=10393120;
RA Hamada Y., Kadokawa Y., Okabe M., Ikawa M., Coleman J.R.,
RA Tsujimoto Y.,
RT "Mutation in ankyrin repeats of the mouse Notch2 gene induces early
RT embryonic lethality."
RL Development 126:3415-3424(1999).
RN [5]
RP DEVELOPMENTAL STAGE, AND ALTERNATIVE SPLICING.
RX MEDLINE=95333693; PubMed=7609614;
RA Higuchi M., Kiyama H., Hayakawa T., Hamada Y., Tsujimoto Y.;
RT "Differential expression of Notch1 and Notch2 in developing and adult
RT mouse brain."
RL Brain Res. Mol. Brain Res. 29:263-272(1995).
RN [6]
RP POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF MET-1699.
RX MEDLINE=21523956; PubMed=11518718;
RA Saxena M.T., Schroeter E.H., Mumm J.S., Kopan R.;
RT "Murine notch homologs (N1-4) undergo presenilin-dependent
RT proteolysis."
RL J. Biol. Chem. 276:40268-40273(2001).
RN [7]
RP POST-TRANSLATIONAL PROCESSING, AND MUTAGENESIS OF MET-1699.
RX MEDLINE=21374376; PubMed=11459941;
RA Mizutani T., Taniguchi Y., Aoki T., Hashimoto N., Honjo T.;
RT "Conservation of the biochemical mechanisms of signal transduction
RT among mammalian Notch family members."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9026-9031(2001).
CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands
CC Jagged1, Jagged2 and Delta to regulate cell-fate determination.
CC Upon ligand activation through the released notch intracellular
CC domain (NICD) it forms a transcriptional activator complex with
CC RBP-J kappa and activates genes of the enhancer of split locus.
CC Affects the implementation of differentiation, proliferation and
CC apoptotic programs (By similarity). May play an essential role in
CC postimplantation development, probably in some aspect of cell
CC specification and/or differentiation.
CC -!- SUBUNIT: Heterodimer of a C-terminal fragment N(TW) and a N-
CC terminal fragment N(EC) which are probably linked by disulfide
CC bonds.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following
CC proteolytical processing NICD is translocated to the nucleus.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=O35516-1; Sequence=Displayed;
CC Name=2;
CC IsoId=O35516-2; Sequence=VSP_001405;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Expressed in the brain, liver, kidney,
CC neuroepithelia, somites, optic vesicles and branchial arches, but
CC not heart.
CC -!- DEVELOPMENTAL STAGE: Expressed in the embryonic ventricular zone,
CC the postnatal ependymal cells, and the choroid plexus throughout
CC embryonic and postnatal development.
CC -!- PTM: Synthesized in the endoplasmic reticulum as an inactive form
CC which is proteolytically cleaved by a furin-like convertase in the
CC trans-Golgi network before it reaches the plasma membrane to yield
CC an active, ligand-accessible form. Cleavage results in a C-
CC terminal fragment N(TW) and a N-terminal fragment N(EC). Following
CC ligand binding, it is cleaved by TNF-alpha converting enzyme
CC (TACE) to yield a membrane-associated intermediate fragment called
CC notch extracellular truncation (NEXT). This fragment is then
CC cleaved by presenilin dependent gamma-secretase to release a
CC notch-derived peptide containing the intracellular domain (NICD)

```

CC from the membrane.
 CC -1- PPM: Phosphorylated.
 CC -1- SIMILARITY: Belongs to the NOTCH family.
 CC -1- SIMILARITY: Contains 35 EGF-like domains.
 CC -1- SIMILARITY: Contains 2 Lin/Notch repeats.
 CC -1- SIMILARITY: Contains 6 ANK repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D32210; BAA22094.1; -
 CC EMBL; X68279; CAA48340.1; -
 CC EMBL; U31881; AAC52924.1; -
 CC PIR; A49175; A49175.
 CC HSP; P16109; IFSB.
 CC MGD; MGI:97364; Notch2.
 CC GO; GO:0005887; C: integral to plasma membrane; IC.
 CC GO; GO:0005515; F: protein binding; IPI.
 CC GO; GO:0002011; P: morphogenesis of an epithelial sheet; IMP.
 CC GO; GO:0007219; P: N signaling pathway; IC.
 CC InterPro; IPR002110; ANK_hydroxyl_S.
 CC InterPro; IPR000152; Asx_hydroxyl_S.
 CC InterPro; IPR000742; EGF_2.
 CC InterPro; IPR001881; EGF_Ca.
 CC InterPro; IPR001438; EGF_II.
 CC InterPro; IPR006209; EGF_Like.
 CC InterPro; IPR02049; Laminin_EGF.
 CC InterPro; IPR08297; Notch_dom.
 CC InterPro; IPR000800; Notch_dom.
 CC Pfam; PF00023; ank; 6.
 CC Pfam; PF00008; EGF; 34.
 CC Pfam; PF00066; notch; 2.
 CC PIRSF; PIRSF002279; Notch; 1.
 CC PRINTS; PR00010; EGFBLCOO.
 CC PRINTS; PR00011; EGFAMININ.
 CC PRINTS; PR01452; NOTCH.
 CC SMART; SM00248; ANK; 6.
 CC SMART; SM00179; EGF_CA; 23.
 CC SMART; SM00004; NL; 3.
 CC PROSITE; PS02097; ANK_REPEAT_REGION; 1.
 CC PROSITE; PS50088; ANK_REPEAT; 4.
 CC PROSITE; PS00010; ASX_HYDROXYL; 22.
 CC PROSITE; PS00022; EGF_1; 33.
 CC PROSITE; PS01186; EGF_2; 27.
 CC PROSITE; PS50026; EGF_3; 35.
 CC PROSITE; PS01187; EGF_CA; 22.
 CC Receptor; Transcription regulation; Activator; Differentiation;
 CC Developmental protein; Repeat; ANK repeat; EGF-like domain;
 CC Transmembrane; Glycoprotein; Signal; Phosphorylation;
 CC Alternative splicing.
 CC SIGNAL 1 25 POTENTIAL.
 CC CHAIN 26 2470 NEUROGENIC LOCUS NOTCH HOMOLOG PROTEIN 2.
 CC CHAIN 1666 2470 NOTCH EXTRACELLULAR TRUNCATION.
 CC CHAIN 1697 2470 NOTCH INTRACELLULAR DOMAIN.
 CC DOMAIN 26 1677 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 1678 1698 POTENTIAL.
 CC DOMAIN 1699 2470 CYTOPLASMIC (POTENTIAL).
 CC DOMAIN 26 63 EGF-LIKE 1.
 CC DOMAIN 64 102 EGF-LIKE 2.
 CC DOMAIN 105 143 EGF-LIKE 3.
 CC DOMAIN 144 180 EGF-LIKE 4.
 CC DOMAIN 182 219 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
 CC DOMAIN 221 256 EGF-LIKE 6, (INCOMPLETE).
 CC DOMAIN 258 294 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
 CC DOMAIN 296 334 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
 CC DOMAIN 336 372 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
 CC DOMAIN 373 411 EGF-LIKE 10.
 CC DOMAIN 413 452 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
 CC FT

FT DOMAIN 454 490 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 492 528 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 530 566 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 568 603 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 605 641 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 643 678 EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 680 716 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 718 753 EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 755 791 EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 793 829 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 831 869 EGF-LIKE 22, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 871 907 EGF-LIKE 23, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 909 945 EGF-LIKE 24, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 947 983 EGF-LIKE 25, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 985 1021 EGF-LIKE 26, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1023 1059 EGF-LIKE 27, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1061 1097 EGF-LIKE 28, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1099 1145 EGF-LIKE 29, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1147 1183 EGF-LIKE 30, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1185 1221 EGF-LIKE 31, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1223 1260 EGF-LIKE 32, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1262 1300 EGF-LIKE 33, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1302 1345 EGF-LIKE 34, CALCIUM-BINDING (POTENTIAL).
 FT DOMAIN 1372 1410 EGF-LIKE 35, CALCIUM-BINDING (POTENTIAL).
 FT REPEAT 1418 1454 LIN/NOTCH 1.
 FT REPEAT 1501 1533 LIN/NOTCH 2.
 FT REPEAT 1535 1569 ANK 1.
 Query Match 2.8%; Score 424; DB 1; Length 2470;
 Best Local Similarity 29.6%; Pred. No. 5.8e-16;
 Matches 105; Conservative 44; Mismatches 146; Indels 60; Gaps 19;
 QY 562 SVDNCPN-CYNGDC---ISG-TCHCFGLFLGPDGCRASCP-----VLCSGNGQY 607
 DB 793 NIDECANPCLNQCTCFDDYSGYTCMLPYTKNCOTVLAPCSNFCENAAVCKEAPNF 852
 QY 608 MKGRCLCHSGWKGAECVDPTNQCIDVACSNHGTCTI-----TGTCLCNPGYKGSCEB--VD 661
 DB 853 ESFSLCAPGQGRCTVDVDECIKSPKMNNGVCHNTQGSYVCEPCPFSGMDEEDIND 912
 QY 662 CMDPTCSGRGVCV---RGECHCFVGGVNGTCEPERATCLDQ-CSGHGTF--LPDTCGLCS 714
 DB 913 CLANPCQNGSCVDHVTFCQHPFGIGKQCDQDNECLSEPCNGKGTCSDYNSVTCT 972
 QY 715 CDPSWTGHC--SIEICA-ADCGGHGVCVG---TCRCEDGMGAAC--DQACHPR-C 764
 DB 973 CPAGFHGVHCENNIDECTESSCFNGGTCVDGINSFSLCPVGTGPFCLHDINECSSNPC 1032
 QY 765 AEHGTCDRG----KCECSPGNGEHC-TIAHYLDVVVKEGCPGLCNGNGCTLDLNGMHC 819
 DB 1033 LNAQTCVDGLGTYRCICPLGVTGNKNCOTLVNLCR-----SPCKNKGTCVQEKARPHC 1095
 QY 820 VCQLGWRGAGCDTSMETACGSDSKNDGDGLVDCMDPDCCQLPLCHINPLCLGSPN 874
 DB 1086 LCPFGWDGAYCDV-LNVSKRAA-----ALQKGVVPEHLQCHSGICINAGN 1129
 RESULT 12
 FBPI_STRPU
 ID FBPI STRPU STANDARD; PRT; 1064 AA.
 AC P10079;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Fibropellin I precursor (Epidermal growth factor-related protein 1)
 DE (UEGF-1).
 GN EGF1.
 OS Strongylocentrotus purpuratus (Purple sea urchin).
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
 OC Echinoida; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
 OC Strongylocentrotus.
 OX NCBI_TaxID=7668;
 RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=90112459; PubMed=25114273;
RA Delgadillo-Reynoso M.G., Rollo D.R., Hursh D.A., Raff R.A.;
RT "Structural analysis of the uEGF gene in the sea urchin
strongylocentrotus purpuratus reveals more similarity to vertebrate
than to invertebrate genes with EGF-like repeats."; J. Mol. Evol. 29:314-327(1989).
RL [2]
RN
RP SEQUENCE OF 279-476 AND 781-1064 FROM N.A.
RX MEDLINE=87319677; PubMed=3498216;
RA Hursh D.A., Andrews M.E., Raff R.A.;
RT "A sea urchin gene encodes a polypeptide homologous to epidermal
growth factor."; Science 237:1487-1490(1987).
RL [3]
RN
RP AVIDIN-LIKE DOMAIN.
RX MEDLINE=89196806; PubMed=2784773;
RA Hunt L.T., Barker W.C.;
RT "Avidin-like domain in an epidermal growth factor homolog from a sea
urchin."; FASEB J. 3:1760-1764(1989).
RL [4]
RN
RP CHARACTERIZATION.
RX MEDLINE=91285254; PubMed=20607114;
RA Bisgrove B.W., Andrews M.E., Raff R.A.;
RT "Fibropellins, products of an EGF repeat-containing gene, form a
unique extracellular matrix structure that surrounds the sea urchin
embryo."; Dev. Biol. 146:89-99(1991).
RL
CC -!- FUNCTION: Form the apical lamina, a component of the extracellular
matrix.
CC
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR. IN VESICLES IN THE CYTOPLASM
OF UNFERTILIZED EGGS, THEN TO THE BASE OF THE HYALIN LAYER
THROUGHOUT DEVELOPMENT AND FINALLY IN THE APICAL LAMINA IN LATE
EMBRYOS AND EARLY LARVAE.
CC
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=IA;
CC IsoId=P10079-1; Sequence=Displayed;
CC Name=IB;
CC IsoId=P10079-2; Sequence=VSP_000451;
CC -!- DEVELOPMENTAL STAGE: Moderate levels in unfertilized eggs and
during early cleavage, then rapidly increases in abundance between
late morula and mesenchyme blastula stages to maximal levels
maintained through subsequent stages. Expressed both maternally
and zygotically.
CC
CC -!- SIMILARITY: Contains 21 EGF-like domains.
CC -!- SIMILARITY: Contains 1 CUB domain.
CC -!- SIMILARITY: THE C-TERMINAL DOMAIN OF THIS PROTEIN IS SIMILAR TO
AVIDIN/STREPTAVIDIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; L08692; AAA62164.1; -;
CC EMBL; L08692; AAA62163.1; -;
CC EMBL; X17530; CAA35571.1; -;
CC EMBL; M17421; AAA30050.1; -;
CC EMBL; X17533; CAA35573.1; -;
CC PIR; A40136; A40136.
CC HSSP; P01132; 15GF.
CC InterPro; IPR001152; Asx_hydroxyl_S.
CC InterPro; IPR005469; Avidin.
CC InterPro; IPR005468; Avidin/str.
CC InterPro; IPR000859; CUB.
CC InterPro; IPR000742; EGF_2.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR001438; EGF_II.

DR InterPro; IPR006209; EGF_like.
DR Pfam; PF01382; Avidin; 1.
DR Pfam; PF00431; CUB; 1.
DR Pfam; PF00008; EGF; 21.
DR PRINTS; PR00709; AVIDIN.
DR PRINTS; PR00010; EGFBL00D.
DR SMART; SM00042; CUB; 1.
DR SMART; SM00175; EGF_Ca; 20.
DR PROSITE; PS00010; ASX_HYDROXYL; 19.
DR PROSITE; PS00577; AVIDIN; 1.
DR PROSITE; PS01180; CUB; 1.
DR PROSITE; PS00022; EGF_1; 19.
DR PROSITE; PS01186; EGF_2; 19.
DR PROSITE; PS00026; EGF_3; 21.
DR PROSITE; PS01187; EGF_Ca; 18.
KW Biotin; Alternative splicing; EGF-like domain; Repeat; Signal;
KW Glycoprotein; Calcium-binding.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1064 FIBROPELLIN I.
FT DOMAIN 20 55 EGF-LIKE 1.
FT DOMAIN 62 175 CUB.
FT DOMAIN 176 212 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 214 250 EGF-LIKE 3, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 252 288 EGF-LIKE 4, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 290 326 EGF-LIKE 5, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 328 364 EGF-LIKE 6, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 366 402 EGF-LIKE 7, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 404 440 EGF-LIKE 8, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 442 478 EGF-LIKE 9, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 480 516 EGF-LIKE 10, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 518 554 EGF-LIKE 11, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 556 592 EGF-LIKE 12, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 594 630 EGF-LIKE 13, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 632 668 EGF-LIKE 14, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 670 706 EGF-LIKE 15, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 708 744 EGF-LIKE 16, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 746 782 EGF-LIKE 17, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 784 820 EGF-LIKE 18, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 822 858 EGF-LIKE 19, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 860 896 EGF-LIKE 20, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 898 934 EGF-LIKE 21, CALCIUM-BINDING (POTENTIAL).
FT DOMAIN 936 1064 AVIDIN-LIKE.
FT DISULFID 23 34 BY SIMILARITY.
FT DISULFID 28 43 BY SIMILARITY.
FT DISULFID 45 54 BY SIMILARITY.
FT DISULFID 180 191 BY SIMILARITY.
FT DISULFID 185 200 BY SIMILARITY.
FT DISULFID 202 211 BY SIMILARITY.
FT DISULFID 218 223 BY SIMILARITY.
FT DISULFID 223 238 BY SIMILARITY.
FT DISULFID 240 249 BY SIMILARITY.
FT DISULFID 256 267 BY SIMILARITY.
FT DISULFID 261 276 BY SIMILARITY.
FT DISULFID 278 287 BY SIMILARITY.
FT DISULFID 294 305 BY SIMILARITY.
FT DISULFID 299 314 BY SIMILARITY.
FT DISULFID 316 325 BY SIMILARITY.
FT DISULFID 332 343 BY SIMILARITY.
FT DISULFID 337 352 BY SIMILARITY.
FT DISULFID 354 363 BY SIMILARITY.
FT DISULFID 370 381 BY SIMILARITY.
FT DISULFID 375 390 BY SIMILARITY.
FT DISULFID 392 401 BY SIMILARITY.
FT DISULFID 408 419 BY SIMILARITY.
FT DISULFID 413 428 BY SIMILARITY.
FT DISULFID 430 439 BY SIMILARITY.
FT DISULFID 446 457 BY SIMILARITY.
FT DISULFID 461 466 BY SIMILARITY.
FT DISULFID 468 477 BY SIMILARITY.
FT DISULFID 484 495 BY SIMILARITY.
FT DISULFID 489 504 BY SIMILARITY.
FT DISULFID 506 515 BY SIMILARITY.
FT DISULFID 522 533 BY SIMILARITY.


```

FT DISULFID 527 BY SIMILARITY.
FT DISULFID 544 BY SIMILARITY.
FT DISULFID 553 BY SIMILARITY.
FT DISULFID 560 BY SIMILARITY.
FT DISULFID 571 BY SIMILARITY.
FT DISULFID 580 BY SIMILARITY.
FT DISULFID 582 BY SIMILARITY.
FT DISULFID 591 BY SIMILARITY.
FT DISULFID 598 BY SIMILARITY.
FT DISULFID 603 BY SIMILARITY.
FT DISULFID 620 BY SIMILARITY.
FT DISULFID 629 BY SIMILARITY.
FT DISULFID 636 BY SIMILARITY.
FT DISULFID 641 BY SIMILARITY.
FT DISULFID 641 BY SIMILARITY.
FT DISULFID 656 BY SIMILARITY.
FT DISULFID 658 BY SIMILARITY.
FT DISULFID 674 BY SIMILARITY.
FT DISULFID 679 BY SIMILARITY.
FT DISULFID 696 BY SIMILARITY.
FT DISULFID 705 BY SIMILARITY.
FT DISULFID 712 BY SIMILARITY.
FT DISULFID 723 BY SIMILARITY.
FT DISULFID 734 BY SIMILARITY.
FT DISULFID 743 BY SIMILARITY.
FT DISULFID 750 BY SIMILARITY.
FT DISULFID 755 BY SIMILARITY.
FT DISULFID 772 BY SIMILARITY.
FT DISULFID 788 BY SIMILARITY.
FT DISULFID 793 BY SIMILARITY.
FT DISULFID 808 BY SIMILARITY.
FT DISULFID 819 BY SIMILARITY.
FT DISULFID 826 BY SIMILARITY.
FT DISULFID 831 BY SIMILARITY.
FT DISULFID 846 BY SIMILARITY.
FT DISULFID 848 BY SIMILARITY.
FT DISULFID 864 BY SIMILARITY.
FT DISULFID 869 BY SIMILARITY.
FT DISULFID 886 BY SIMILARITY.
FT DISULFID 902 BY SIMILARITY.
FT DISULFID 907 BY SIMILARITY.
FT DISULFID 924 BY SIMILARITY.
FT CARBOHYD 30 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 851 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 477 Missing (in isoform IB).
FT FT 279 /Frid=VSP 000451.
FT CONFLICT 279 L -> S (IN REF. 2).
FT SEQUENCE 1064 AA; 112072 MW; 2E569CA012ED09 CRC64;

Query Match 2.8%; Score 423.5; DB 1; Length 1064;
Best Local Similarity 27.6%; Pred. No. 1.6e-16;
Matches 131; Conservative 59; Mismatches 150; Indels 135; Gaps 30;

QY 545 NDGKESVWS-FLTATE-----SVDCPSN-CYNGDCI-----SGTCHCFLGLG 589
DB 605 NGGACMDVNGFVCTCLPGHEGTNCEINTECASSPCWGLCVQDQNSVCFCLPFTG 664
QY 590 PDGR-----ASCFVCSGNGQYM-----KGRCLCHGKGAECVPTNOCIDVACSNHGT 640
DB 665 IHCGTEIDECASSP--CLNGGQCIDRVDSEVCVCAAGYTAVCQINIDECASAPCQNGV 722
QY 641 CITG-----TICNPGYKESCE-EVD-CMDPTCSGRGVCR-----GECHCFVWGSGTNC 590
DB 723 CVDGNGVYVNCAPGYTGDMNCETIEDBCASMPCLNGACIEMVNGYTCQCVAGYTVGICE 782
QY 691 TPRATCLDQSG-----HGTELPDT---GLSCDPSMTGHDG---SIFICAAD-CGGHGVCV 740
DB 783 TD-----IDECASAPCQNGGVCTDTINGYICACVPGGTGTNCEINIDECASDPLNGICV 838
QY 741 GG-----TCREDGWGAACP--ORACHPR--CAEHGTC-----RGKCECPGNGHECTTA 789
DB 839 DGVNGFVCQCPNYSYGYTCEISLDACRMPCCQNGATCVNMGADYCEVPGVAGQCEID 898
QY 790 HYLDVVKGGCPGL-CNNGRCRTLDLNGHVCQGLWRGAGCDTSMETA-CGDSKNDGD 847
DB 899 -----INE-CASIPQNGGLCIDGTAGYTCQRELGYGVNCE---EVGFC-----DLEG 944
QY 848 GLVDCMDPDCCLOPLCHINPLCLGSPNPLDIIQETQVPVSVQNLHGFYDRKPLVGRDST 907
DB 945 WYNECND-----VTIT-----QVTIT-----KTST 960
QY 908 HIIPGE-----NPFDDGHCACVIRQVMTSDGTPLVGVNISVNNPLFGYITSRQDG 958

DB Db 961 GMLGDYMTYNERALGYAA-----PTVVVGYASNNYDFPSFGFTVVRDNG 1005

RESULT 13
NTCL RAT STANDARD; PRT; 2531 AA.
AC Q07008;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neurogenic locus notch homolog protein 1 precursor (Notch 1).
GN NOTCH1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]_TaxID=10116;
RP SEQUENCE FROM N.A.
RC TISSUE=Schwann Cell;
RX MEDLINE=92111383; PubMed=1764995;
RA Weinmaster G., Roberts V.J., Lemke G.;
RT "A homolog of Drosophila Notch expressed during mammalian
RL development.";
RN [2] Development 113:199-205(1991).
RP REVISIONS TO 1652-1653.
RA Weinmaster G.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP FUNCTION.
RX MEDLINE=21094508; PubMed=11182080;
RA Tanigaki K., Nogaki F., Takahashi J., Tashiro K., Kurooka H.,
RA Honjo T.;
RT "Notch1 and Notch3 instructively restrict bFGF-responsive multipotent
RL neural progenitor cells to an astroglial fate.";
RN [4] Neuron 29:45-55(2001).
RP TISSUE SPECIFICITY.
RX MEDLINE=93202015; PubMed=1295745;
RA Weinmaster G., Roberts V.J., Lemke G.;
RT "Notch2: a second mammalian Notch gene.";
RN [5] Development 116:931-941(1992).
RP TISSUE SPECIFICITY.
RX MEDLINE=21331789; PubMed=11438922;
RA Irvin D.K., Zurcher S.D., Nguyen T., Weinmaster G., Kornblum H.I.;
RT "Expression patterns of Notch1, Notch2, and Notch3 suggest multiple
RL functional roles for the Notch-DSL signaling system during brain
development.";
RN [6] J. Comp. Neurol. 436:167-181(2001).
CC -!- FUNCTION: Functions as a receptor for membrane-bound ligands
CC Jagged1, Jagged2 and Deltal to regulate cell-fate determination.
CC Upon ligand activation through the released notch intracellular
CC domain (NICD) it forms a transcriptional activator complex with
CC RBP-J kappa and activates genes of the enhancer of split locus.
CC Affects the implementation of differentiation, proliferation and
CC apoptotic programs (By similarity). Acts instructively to control
CC the cell fate determination of CNS multipotent progenitor cells,
CC resulting in astroglial induction and neuron/oligodendrocyte
CC suppression.
CC -!- SUBUNIT: Heterodimer of a C-terminal fragment N(TM) and a N-
CC terminal fragment N(EC) which are probably linked by disulfide
CC bonds (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. Following
CC proteolytical processing NICD is translocated to the nucleus (By
CC similarity).
CC -!- TISSUE SPECIFICITY: Expressed in the brain, kidney and spleen.
CC Expressed in postnatal central nervous system (CNS) germinal zones
CC and, in early postnatal life, within numerous cells throughout the
CC CNS. Found in both subventricular and ventricular germinal zones.
CC -!- DEVELOPMENTAL STAGE: In the embryo, highest levels occur between
CC days 12 and 14 and decrease rapidly to much lower levels in the
```

DOMAIN	295	333	EGF-LIKE 8,	CALCIUM-BINDING	(POTENTIAL)
DOMAIN	335	371	EGF-LIKE 9,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	372	410	EGF-LIKE 10,		
DOMAIN	412	450	EGF-LIKE 11,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	452	488	EGF-LIKE 12,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	490	526	EGF-LIKE 13,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	528	564	EGF-LIKE 14,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	566	601	EGF-LIKE 15,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	603	639	EGF-LIKE 16,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	641	676	EGF-LIKE 17,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	678	714	EGF-LIKE 18,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	716	751	EGF-LIKE 19,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	753	789	EGF-LIKE 20,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	791	827	EGF-LIKE 21,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	829	867	EGF-LIKE 22,		
DOMAIN	869	905	EGF-LIKE 23,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	907	943	EGF-LIKE 24,		
DOMAIN	945	981	EGF-LIKE 25,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	983	1019	EGF-LIKE 26,		
DOMAIN	1021	1057	EGF-LIKE 27,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	1059	1095	EGF-LIKE 28,		
DOMAIN	1097	1143	EGF-LIKE 29,		
DOMAIN	1145	1181	EGF-LIKE 30,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	1183	1219	EGF-LIKE 31,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	1221	1265	EGF-LIKE 32,	CALCIUM-BINDING <th>(POTENTIAL)</th>	(POTENTIAL)
DOMAIN	1267	1305	EGF-LIKE 33,		
DOMAIN	1307	1346	EGF-LIKE 34,		
DOMAIN	1348	1384	EGF-LIKE 35,		
DOMAIN	1387	1426	EGF-LIKE 36,		
REPEAT	1445	1480	LIN/NOTCH 1,		
REPEAT	1481	1522	LIN/NOTCH 2,		
REPEAT	1523	1562	LIN/NOTCH 3,		
REPEAT	1917	1946	ANK 1,		
REPEAT	1950	1980	ANK 2,		
REPEAT	1984	2013	ANK 3,		
REPEAT	2017	2046	ANK 4,		
REPEAT	2050	2079	ANK 5,		
REPEAT	1730	1733	POLY-ALA,		
DOMAIN	1891	1894	POLY-GLU,		
DOMAIN	2258	2261	POLY-PRO,		
DOMAIN	2497	2500	POLY-SER,		
SITE	1654	1655	CLEAVAGE BY (FURIN-LIKE PROTEASE) (BY SIMILARITY),		
DISULFID	24	37	BY SIMILARITY,		
DISULFID	31	46	BY SIMILARITY,		
DISULFID	48	57	BY SIMILARITY,		
DISULFID	63	74	BY SIMILARITY,		
DISULFID	68	87	BY SIMILARITY,		
DISULFID	89	98	BY SIMILARITY,		
DISULFID	106	117	BY SIMILARITY,		
DISULFID	111	127	BY SIMILARITY,		
DISULFID	129	138	BY SIMILARITY,		
DISULFID	144	155	BY SIMILARITY,		
DISULFID	149	164	BY SIMILARITY,		
DISULFID	166	175	BY SIMILARITY,		
DISULFID	182	195	BY SIMILARITY,		
DISULFID	189	204	BY SIMILARITY,		
DISULFID	206	215	BY SIMILARITY,		
DISULFID	222	233	BY SIMILARITY,		
DISULFID	227	243	BY SIMILARITY,		

Query Match 2.8%; Score 422.5; DB 1; Length 2531;

Best Local Similarity 30.2%; Pred. No. 7.3e-16;

Matches 108; Conservative 40; Mismatches 139; Indels 71; Gaps 19;

562 SYDNCPNS-CYGNCGDISG-----TCHCFGLFGLPGCR-----ASCVLCSGNGQYMKG- 610

490 NTDECASSPCLHNGKCVDKINEFLCQPKPFGHLCOYDVDECASTP--CWNGAKCLDGP 547

611 ---RCLCHSGWKGAECQVPTNQCIDVACSNHGTCITG---TCICNPGYKGESCEE--VD 661

548 NTYTCTVCTEGYTGTHCEVDIDCEDDDPC-HGLCKDQGVATFTCLCQPGYTGHHCTNINE 606

FT	DOMAIN	570	605	EGF-LIKE 15,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	607	643	EGF-LIKE 16,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	645	680	EGF-LIKE 17,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	682	718	EGF-LIKE 18,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	720	755	EGF-LIKE 19,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	757	793	EGF-LIKE 20,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	795	831	EGF-LIKE 21,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	833	871	EGF-LIKE 22,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	873	909	EGF-LIKE 23,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	911	947	EGF-LIKE 24,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	949	985	EGF-LIKE 25,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	987	1023	EGF-LIKE 26,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	1025	1061	EGF-LIKE 27,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	1063	1099	EGF-LIKE 28,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	1101	1147	EGF-LIKE 29,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	1149	1185	EGF-LIKE 30,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	1187	1223	EGF-LIKE 31,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	1225	1262	EGF-LIKE 32,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	1264	1302	EGF-LIKE 33,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	1304	1343	EGF-LIKE 34,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	1374	1412	EGF-LIKE 35,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	1645	1648	POLY-ALA,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	1994	1997	POLY-LEU,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	2426	2429	POLY-SER,	CALCIUM-BINDING	(POTENTIAL).
FT	DOMAIN	2446	2451	POLY-GLY,	CALCIUM-BINDING	(POTENTIAL).
FT	REPEAT	1420	1456	LIN/NOTCH 1,	CALCIUM-BINDING	(POTENTIAL).
FT	REPEAT	1503	1535	LIN/NOTCH 2,	CALCIUM-BINDING	(POTENTIAL).
FT	REPEAT	1827	1871	ANK 1,	CALCIUM-BINDING	(POTENTIAL).
FT	REPEAT	1876	1905	ANK 2,	CALCIUM-BINDING	(POTENTIAL).
FT	REPEAT	1909	1939	ANK 3,	CALCIUM-BINDING	(POTENTIAL).
FT	REPEAT	1943	1972	ANK 4,	CALCIUM-BINDING	(POTENTIAL).
FT	REPEAT	1976	2005	ANK 5,	CALCIUM-BINDING	(POTENTIAL).
FT	REPEAT	2009	2038	ANK 6,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	28	41	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	35	51	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	53	62	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	68	79	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	73	90	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	92	101	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	109	121	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	115	131	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	133	142	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	148	159	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	153	168	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	170	179	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	186	198	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	192	207	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	209	218	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	225	236	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	230	246	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	248	257	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	264	275	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	269	284	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	286	295	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	302	315	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	309	324	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	326	335	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	342	353	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	347	362	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	364	373	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	379	390	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	384	401	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	403	412	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	419	433	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	427	442	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	444	453	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	460	471	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	465	480	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	482	491	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	498	509	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	503	518	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	520	529	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).
FT	DISULFID	536	547	BY SIMILARITY,	CALCIUM-BINDING	(POTENTIAL).

Query Match 2.8%; Score 420; DB 1; Length 2471;
 Best Local Similarity 29.4%; Pred. No. 9.9e-16;
 Matches 104; Conservative 44; Mismatches 148; Indels 58; Gaps 19;

QY	562	SVDNCPNS-CYGNQDC---ISG-TCHCFGLPGDCGRASCP-----VLCSGNGQY 607
DB	795	NIDECASNPLNQGTCLDDVSGYTCWMLPYTGKNCQTVLAPCSNPNCENAAVCKEAPNF 854
QY	608	MKGRCLCHSGKGAECVPTNQCIDVACSNHGTG---ITGT---CICNPGYKGCSEB---VD 661
DB	855	ESFTCLCAPGQGORCTVDDECVCSPKPMNNGICHTQGSYMCBCPPGSGMDCBEDIND 914
QY	662	CMDPTCSGRGVGV---RGECHCFVGMGGTNCETPRATCLDQ-CSGHGTF--LPDTGLCS 714
DB	915	CLANPCQNGGSCVDKVNFTFSCLLCPGFVGDRCQTDMECLSEPCCKNGGTCSDYVNSYTC 974
QY	715	CDPSWTGHDC--STEICA-ADCGHGVGVGG---TCRCEDGWMGAAC--DQRACHPR-C 764
DB	975	CPAGFHGVHCENNIDECTESSCFNGGTCVDGINSFSLCPVGTGTPGPFCLHDHINECSSNPC 1034
QY	765	AHSGTCRDG---KCECSPGMNGEHCTIAHLYDRVVKEGCPGLGNGRCCTLDLNGWHCV 820
DB	1035	LSNGFTCVDLGTYRCTPLGYTGKVC-----QTLVNLCSPPCKNKGTCQAERARPC 1088
QY	821	COLGWRGAGCDTSMETACGSKNDGDLVDCMDPDCCLQPLCHINPCLGSPN 874
DB	1089	CPGWDGAYCDV-LNVSKAA-----ALQKGVPEVHLCHQHSIGICINAGN 1131

RESULT 15
 JAG1_RAT
 ID JAG1_RAT STANDARD; PRT; 1219 AA.
 AC Q63722; P70640;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Jagged 1 precursor (Jagged1).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN (1)_TaxID=10116;
 RP SEQUENCE FROM N.A.
 RC TISSUE=Sciatic nerve;
 RX MEDLINE=95211842; PubMed=7697721;
 RA Lindsell C.E., Shawber C.J., Boulter J., Weinmaster G.;
 RT "Jagged: a mammalian ligand that activates Notch1.";
 RL Cell 80:909-917(1995).
 CC -!- FUNCTION: Ligand for multiple Notch receptors and involved in the
 CC mediation of Notch signaling. May be involved in cell-fate
 CC decisions during hematopoiesis. Enhances fibroblast growth
 CC factor-induced angiogenesis (in vitro). Seems to be involved in
 CC early and late stages of mammalian cardiovascular development.
 CC Inhibits myoblast differentiation. May regulate fibroblast
 CC growth factor-induced angiogenesis.
 CC -!- SUBUNIT: Interacts with NOTCH1, NOTCH2 AND NOTCH3 (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Widely expressed in a variety of tissues.
 CC -!- DEVELOPMENTAL STAGE: Expression is seen in E11.5-E14.5 embryos in
 CC four distinct regions of the ventricular zone in the developing
 CC spinal cord.
 CC -!- SIMILARITY: Contains 15 EGF-like domains.
 CC -!- SIMILARITY: Contains 1 DSL domain.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC


```

QY 942 -----SFVN-----NPLFGYTIHQDGSFDLVT--NGGISILLRPERA 977
Db 669 NDCSQNPCHYGGTCRDLVNDPYCDKXGKWKTKCHSRDQCDEATCNGG----- 718
QY 978 PFITQBHTLWLPDRFPVMEIIMRHEHEIPSC-----DLSNFAR-----PNP----- 1021
Db 719 -----TCVDEVTFCM-----CPGWEGTTCNIARNSSCLPNPCHNGG 757
QY 1022 --VVSPSPLTSFASCAE--KGPVPEIQALQEEISISGCKWLSYLSR----- 1067
Db 758 TCWVNGD---SFTCVCKEGEGPCTQ-----NTNDCSPHCYNSGTCTVDGDNWYRC 806
QY 1068 --TPGVKSVLRISLTHPTIPFNLMKVHLMVAVEGRLEFRKWFAAAPDLISYFIWDKTDVYN 1125
Db 807 ECAPGFAG-----PDCRININECQSSPCAFGA-----TCVDEINGY----- 842
QY 1126 QKVFGLSEAFVSVGYEYESCPDLILWEKRTTVLQGYEI-DASKLGWSLDKXHALNIQSG 1184
Db 843 -----QCICPPGHSGAKCHEV---SGRSCITMGRVILDGAK---WDDDCNTCQCINGR 889
QY 1185 I-----LHKGNGENQFVSQQPPVIGS-----IMGNRRRRSISPCSCNGLAD 1225
Db 890 VACSKVWCGRPCRELKHGHECPNGQSCIPVLDQCFVRPCTGAGBCRSS----- 940
QY 1226 GNKLAPVALTCGSDGSLYVD-----FNYIRIFPSGNVTNIL--ELRKNDFRHSHP 1277
Db 941 ---LQPVKTCTSDS--YYQDNCANITFTFNKEMWSPGLTTEHICSELRLNLIKXVSA 994
QY 1278 AHKYVLATDPMGAVFLSDNSRVPFKISTVVVKDIAKNSRVVAGTGDCQLPDDTRCG 1337
Db 995 EYSTIACEPS-----LSANN-----EIHVAISAEDIRDDGNPVKEITKIIDLVSKR-- 1042
QY 1338 DGGKATEATLTNPR 1351
Db 1043 DGNSSLIAVAEVR 1056

```

Search completed: June 24, 2004, 16:19:49
Job time : 31 secs

This Page Blank (uspto)